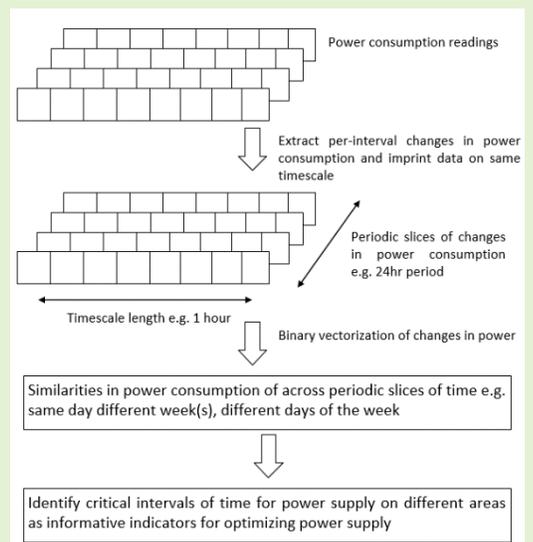


A Sensory Similarities Approach to Load Disaggregation of Charging Stations in Internet of Electric Vehicles

Qi Liu, *Senior Member, IEEE*, Kondwani M. Kamoto, Xiaodong Liu, *Senior Member, IEEE*, Yonghong Zhang, Zaiqiang Yang, Mohammad R. Khosravi, Yanwei Xu and Lianyong Qi

Abstract—Intelligent transportation systems (ITSs) have become popular in recent years as an essential requirement for safer and more efficient transportation systems. Internet of Electric vehicles (IoEV) as well as their hybrid forms provide an ideal means of supporting sustainability within an ITS. The control of charging/discharging of EV is still a challenge, despite the tremendous research progress to date in the field. In this paper, the use of charging station data and binary vectorization are proposed in order to provide timely insights on the dynamic behavior of charging processes. A Bag-of-Power-States model has been created for similarity measurement of charging stations within given time periods. The results of experimentations using synthetic data have shown that the proposed *Bag-of-Power-States* model is computationally feasible and provides useful results for optimizing the scheduling of power supply to charging stations that may be located across a wide range of distances, over the same period of time.

Index Terms—Sensory Similarities, Bag-of-Power-States, Internet of Electric Vehicles, Intelligent Transportation Systems



I. Introduction

Automation along with the quantity and variety of technologies within vehicles and the transportation industry have constantly evolved over time in recent years, with the main driving forces being to ensure utmost safety and efficiency of transportation services [1].

Advancements in sensing technologies, communication technologies, and control technologies have resulted in a more modern and connected landscape for transportation services, resulting in the establishment of Intelligent Transportation Systems (ITSs) [2-4].

While the benefits of ITS cannot be understated with Advanced Driver-Assistance Systems (ADAS), Cooperative Adaptive Cruise Control (CACC), Lane Departure Warning/Keeping (LDW/LDK), and Collision Avoidance Systems (CAS) helping to prevent road accidents and save lives, plus efficient routing and traffic management strategies

saving time and money, technological advancements present new forms of challenges for the automotive industry [5]. These challenges include unprecedented levels of data that needs to be processed in a timely matter in order to be acted upon, ensuring optimal performance of the services, particularly communications, while accounting for varying vehicle motions and a variety of locations, and providing adequate measures of security and privacy of data and infrastructure at all levels both within and beyond a vehicle. The use of established and new techniques in big data analytics and deep learning models have sought to address the increasing amounts of heterogeneous data in an ITS as well as improving vehicle control, with the introduction of Intrusion Detection Systems (IDSs) into a vehicular network [6] and/or improvements in a Vehicular Ad-Hoc Network (VANET) addressing challenges in security issues, privacy leakage and trustworthy threats [7].

Electric vehicles (EVs) have emerged as an essential component for sustainable transportation systems, with

Q. Liu is with the Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAET), Nanjing University of Information Science & Technology, No. 219, Ningliu Road, Nanjing, 210044, China and Shandong Beiming Medical Technology Ltd, Jinan 250000, China (e-mail: qi.liu@nuist.edu.cn).

K. M. Kamoto is with the School of Computer and Software, Nanjing University of Information Science & Technology, No. 219, Ningliu Road, Nanjing, 210044, China (e-mail: k_kamoto@yahoo.co.uk).

X. Liu is with the School of Computing, Edinburgh Napier University, 10 Colinton Road, Edinburgh, EH10 5DT, UK (e-mail: x.liu@napier.ac.uk).

Y. Zhang is with the School of Automation, Nanjing University of Information Science & Technology, Nanjing, Jiangsu 210044, China (zyh@nuist.edu.cn).

Z. Yang is with the Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science & Technology, Nanjing, China (yzq@nuist.edu.cn)

M. R. Khosravi is with the Department of Computer Engineering, Persian Gulf University, Bushehr, Iran. (e-mail: m.khosravi@mehr.pgu.ac.ir).

Y. Xu is with the School of Software, Tianjin University, China (e-mail: xuyanwei@tju.edu.cn).

L. Qi is with the School of Information Science and Engineering, Qufu Normal University, China (e-mail: lianyongqi@qfnu.edu.cn).

increasing rates of adoption across the globe and a number of notable implementations in US and European countries [8-10]. The increasing rates of adoption can be attributed to the establishment and continual improvement of global standards and practices, lowering costs of EV batteries, and reduced complexity of charging schemes, all of which has resulted in reduced anxiety regarding EVs and their related operations [11].

This paper aims to provide an approach to addressing sustainability in an ITS by optimizing supply of power to charging infrastructure in accordance with the dynamic nature of the environment within transportation systems, as shown in Fig. 1. The research work presented here investigates how the use of power consumption obtained from different forms of charging substations can provide information on how the consumption evolves over time which can be put to use to better coordinate delivery of adequate power supply for different locations in a timely manner. The proposed method makes use of evolving changes in electric power consumption and allows for comparison of charging behavior in substations for varying intervals of time within a single location, or amongst a set of locations. Major contribution of the proposed model is to extend the work in a sustainable ITS based on the works established in the field of load monitoring, which studies how to categorize power consumption among charging stations. While the main body of work in this field of study has primarily focused on identification of appliance and/device power consumption it is the authors' belief that such an analysis of the power consumption can be extended to ITS infrastructure, which includes roadside infrastructure and road administration infrastructure. Binary encoding of the data is considered in this work as the primary means of efficient comparison given the high volume of data that will be required to be processed as is

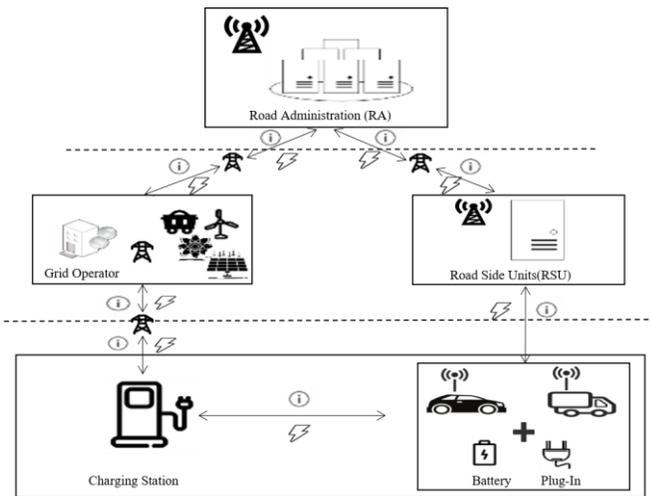


Fig. 1. Sustainable and Dynamic Nature of EVs in an ITS

typical in current ITSs.

The rest of this paper is organized with following sections. Related solutions and research efforts have been discussed in Section II, followed by a preliminary section on the concept of Operational Taxonomic Units (OTUs), which are the basis for the definition of binary similarity and distance measures. Section IV presents the proposed research work which includes the proposed binary encoding scheme as well as the means for quantifying similarities in power consumption across various

charging substations. Section V provides details on the experiments used to evaluate the proposed similarity model, and summarizes the findings from the experimentation. Finally, the conclusion of the paper as well as potential work to be conducted in the future is summarized in Section VI.

II. RELATED WORK

A. Overview of Intelligent Transportation Systems

Improvements to sensor and communications technologies have led to enhanced collection practices with data sources including GPS, video cameras, sensors, passive sources i.e. mobile, internet, and smart cards used for vehicle services [12]. Data analytics using Linear Regression [13], Artificial Neural Networks (ANNs) [14], Decision Trees (DTs) [15], and k-means [16] have shown promising results for road traffic analysis and flow prediction, planning of public transportation and control, personal travel routing services, and asset maintenance.

Deep Learning (DL) models were also increasingly deployed to analyze the heterogeneous data and provide meaningful information as well viable control actions to be performed by people and vehicles [17]. Deep learning has seen wide adoption in recent years due to its good performance for solving complex non-linear control problems, and good generalization performance of deep learning models to new scenarios. Different DL-based models (e.g. CNN, RNN, Reinforcement Learning, etc.) have shown good performance for various problems including vision recognition, image processing and motion detection [18-20]. Despite the current success of big data analytics and deep learning the dynamic nature of transportation system poses significant hurdles that still need to be overcome. Data privacy, storage, and processing are still largely unsolved in terms of real-time practicality, with deep learning models requiring significant amounts of data that covers a vast number of scenarios in order to be useful for practical and also being resilient to adversarial attacks.

Reinforcement learning has greatly simplified verification and validation of performance of algorithms and models for ITS [21-23]; however, there are challenges in the definition of a suitable training environment in accordance with real world parameters as well as the definition of meaningful reward functions for taking multiple actions in a given training, which are typical scenarios of ITSs (e.g. slow the car down while turning to the left). Therefore, improvement to sensors, devices and their communication have directly led to a more sustainable system due to high energy efficiency and increased longevity of components, particularly in a smarter environment with a wider range of heterogenous devices and means of communication.

While sensing, communications, and control technologies have become foundational elements within ITS systems, Electric Power and Energy System (EPES) share a similarly important role for ITS [24]. Modern EPESs by nature support bidirectional flow of both electric power and data, allowing more efficient and effective practices for perception and control. Harnessing the power consumption data from ITS infrastructure presents the means of optimizing power supply in order to support ITS operations across various locations.

Furthermore, the introduction of promising Micro-Grid architectures (MGs) [25, 26], which can utilize Renewable Energy Sources (RESs) that are locally available can enable resilience and also geographically extend current ITS operations.

B. Overview of Electric Vehicles(EV) and Vehicle-to-Grid (V2G)

EVs, as well as Plug-in EVs (PEVs) have become key components of a sustainable ITS, with potential benefits for the EPESs through exchange of power using Vehicle-to-Grid (V2G) interactions [27]. While V2G is the most promising use of EVs, it has currently met with a number of significant technical and operational shortcomings, limiting its use in practice [28]. V2G interactions are therefore not employed and discussed in the work of this paper.

Two-stage stochastic programming which is an approach used to model problems that involves uncertainty, has been recently shown to be a useful technique for solving the scheduling problem using the centralized approach [29, 30]. Following this approach, corresponding scheduling problems were examined and optimized in the context of a renewable energy microgrid [31], where the authors demonstrated how uncertainties such as EV connectivity could be handled by transforming the mathematical model in order to account for the updated problem inputs. In addition, Markov Decision Process (MDP) [32] and multi-agent systems [33] have been used to address the challenges of decentralized scheduling and the stochastic nature of EV charging patterns.

Supervised and unsupervised learning/classification, e.g. K-means [34], Self-Organizing Map (SOM) [35], Markov models [36], Support Vector Regression (SVR) [37] and Decision Tree (DT) [38] have been the basis for a majority of research approaches to solving clustering of EV load profiles using both supervised learning and unsupervised learning. A defining factor of these approaches is the complexity of the data that the models are trained, ranging from temporal, spatial datasets to spatial-temporal datasets. Present research work shows that clustering techniques can offer efficient approaches to solving the locating challenges of charging stations, and their effective expansion. However, different solutions may suffer from diverse charging/loading features at clusters, as such major obstacle still needs to be overcome.

Dynamic programming (DP) has been a key general baseline approach for solving problems related to EV charging (and EV tasks in general), with the charging activities of EVs being restated in the form of an optimization problem. The use of DP and historical data was explored and discussed in order to produce data-driven models that can make optimal EV charging decisions in real-time [39]. The authors evaluated models based on k-Nearest Neighbors (k-NN), Shall Neural Networks (SNNs), and Deep Neural Networks (DNNs), which were trained on optimal decisions derived from applying DP to the historical data. The outcomes of experimentation demonstrated that such a combination of approaches can lead to significant reduction in charging costs due to efficient and effective EV charging, with combination of the DP and DNN showing the

most promising results. Accordingly, the DP-based research work has shown promising results; however, major hurdles still exist in the formulation of optimization processes as more variables are introduced into EV charging scenarios.

Data-driven approaches have recently highlighted the importance of scheduling and control decisions on wide range of EV charging data. Similarity of a number of charging activities was examined from a variety of regions in Nanjing, China [40], where the analysis of EV time-energy characteristics and charging demand was conducted using Kernel Density Estimation (KDE). In this paper, the authors also explained the relationship of charging parameters derived from operational data, which can be employed to arrange the distribution of Renewable Energy Sources (RES) in a smart grid with rational guidance. The detection of domestic charging PEVs was presented in [41], with data-driven classification models, e.g. k-NN obtained with high frequency power data. Experimentation has demonstrated that a data-driven approach can result in high levels of accuracy (80-90%) when it comes to identify households that use PEVs. In addition, two different datasets were examined in [42], where one set of data was obtained from customer profiles (charging records), and the other derived from outlet measurements (station data). Four models were employed to analyze the datasets and perform the prediction, i.e. Time Weighted Dot Product based Nearest Neighbor (TWDP-NN), Modified Pattern Sequence Forecasting (MPSF), SVR and Random Forest. The comparison was conducted with regards to speed, accuracy, and privacy concerns.

While steady progress is being made to ensure that the necessary legal, business, and production aspects are in place, there are still quite a number of challenges to be overcome [43]. Recent research activities have focused on the use of various planning, and predicting strategies to recognize potential charging features of EVs. However, the unpredictability of EV charging/discharging patterns still presents a challenge in the field of study, particularly when consideration the time-sensitive nature of the decision making process for these processes. Therefore, the focus of this paper is to make a valuable contribution towards solving this problem by providing insight in the dynamic patterns within the processes.

III. PRELIMINARY WORK

Binary vectors provide meaningful representations of data and allow for easier comparisons amongst two or more objects or patterns [44]. Similarity/dissimilarity can therefore be measured and used for classification and clustering of data. Numerous binary similarity measurement approaches have been conceived in order to more accurately analyze data in various fields including ecology, biology, ethnology, taxonomy, image analysis, and text analysis.

The binary encoding process of objects varies depending on the approach or model architecture that is adopted in order to define binary embedding. In this paper, the binary embedding is defined by breaking down load profiles discovered in given periods of time into their constituent states, termed power

TABLE I
OTU EXPRESSIONS [45]

$j \backslash i$	1 (Presence)	0 (Absence)	Sum
1 (Presence)	$a = i \odot j$	$b = \bar{i} \odot j$	$a + b$
0 (Absence)	$c = i \odot \bar{j}$	$d = \bar{i} \odot \bar{j}$	$c + d$
Sum	$a + c$	$b + d$	$n = a + b + c + d$

states. The union of the power states is then used as the basis for binary encoding, which includes the processes of denoting a power state (presence/absence) in a given period of energy usage

Operational Taxonomic Unit (OTU) [45] is used to measure binary similarity, i.e. the distance between objects. It can be expressed using a 2x2 contingency table similar to Table I, where i and j are objects or patterns that are under comparison and represented in the form of binary feature vectors. n is the number of features or dimensions.

The OTU expressions are defined as follows:

- a represents the *positive matches* which is the count of features that have values of 1 for both i and j ;
- b represents the *i presence mismatch* which is the count of values that i and j are (1, 0);
- c represents the *i absence mismatch* which is the count of values that i and j are (0, 1);
- d represents the *negative matches* which is the count of features that have values of 0 for both i and j .

Accordingly, the diagonal sum $a + d$ is the total number of matches (positive and negative) of both i and j ; whereas the diagonal sum $b + c$ is the number of mismatches (presence/absence) of both i and j . The total sum through both expressions are the same as the total number of features or dimensions, i.e. n .

The OTU and its binary feature/dimension expressions can enable more information to be extracted from aggregated power data than would normally be possible. This is mainly due to providing the capabilities to discover similarities in energy usage between two given periods of time. Furthermore, the absence and/or presence of certain power states can assist in identifying which load profiles are common to the given time periods and which ones are unique to a specific time period.

IV. PROPOSED WORK

In this section, the analysis of charging power of EVs in an ITS is conducted, including initial steps to capture and transmit the data, and then additional processing steps in order to discover patterns of usage. The discovered patterns are then used as or fed into further processes such as load classification, forecasting, demand response programs, etc. The charging power (i.e. the consumption data) via the charging substations is processed and then used to dynamically quantify similarities in the power consumption across the various components of the infrastructure. In this paper, a similarity measurement approach is proposed, where the workflow is visualized in Fig. 2. A detailed explanation of the steps within the workflow is provided in the following subsections.

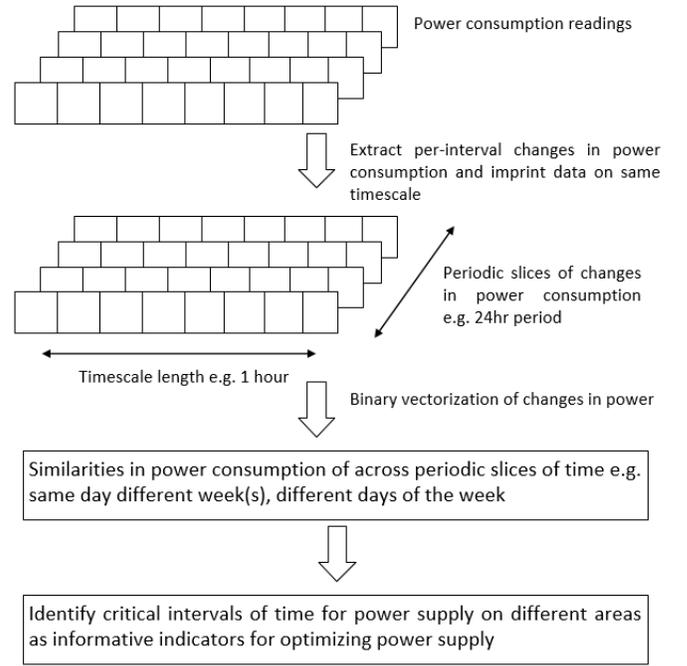


Fig. 2. Workflow of the Proposed Work

A. Power Data Acquisition

Capturing and transmitting the energy data poses a number of challenges, with one of the most challenging aspects being ensuring that the data is useful enough for analysis and future processes. This work proposes a data encoding method which can assist with this challenging task. Electric power data is transmitted at varying frequencies based on the capabilities of the metering technologies used to capture it, this makes utilizing data received from various sources at varying frequencies challenging work. To address this we propose transforming the data into slices representative of periodic intervals e.g. 10 minutes, 1 hour, using a default sampling frequency. Such a data transformation enables easier interpretation of the obtained readings at varying levels of frequency, and can assist in verifying suitable sampling frequency across a set of infrastructure components as well as missing values in the data. This work proposes a default periodic interval of 1 hour, and sampling frequency of either 1 second or 1 minute in order the best capture the varying patterns in power consumption across time with an adequate level of granularity. Both the periodic interval and sampling frequency can be adjusted accordingly, depending on the requirements of future processes.

B. A Bag-of-Power-States Model

The proposed model makes use of power states in order to determine similarity among two or more periods of energy usage. Each period of power usage is defined according to the patterns of power consumption with higher consumption levels indicating more burden on the infrastructure due to increase in activity, and vice versa. Non-Intrusive Load Monitoring (NILM) [46] based procedures have been employed in order to facilitate feature extraction. Furthermore, interval-based changes in power have been manipulated to denote activity in a

given aggregated electric power signal. The following subsections provide an overview of the individual steps involved in the model.

1) Feature Extraction

The slices of power data produced by the data are first examined in order to determine the power usage, with the slices spanning a single day set as the typical window for examining the power consumption. The changes in power between each interval in a slice are used as features that can be used to determine possible activity. This is typically done by extracting the magnitude-wise changes at each interval as shown in (1).

$$\Delta P_t = P_{t+1} - P_t \quad (1)$$

where ΔP_t is the change in power between time intervals $t + 1$ and t . The feature extraction process results in a vector which contains the individual changes in power that have been discovered power data obtained from a single source.

Given that electric power can contain noise due to transmission and interference, and not all the features in power are an indicator of activity, it is important that only the useful features are kept. This process is carried out by first establishing a feature exclusion threshold $threshold_f$ and setting any values of p that are below the assigned threshold value to zero. This results in a somewhat sparse vector depending on the features that are retained.

2) Modeling Power Usage

The next stage involves defining the power usage in the form of power usage models. The models consist of two or more states, and each state only retains the magnitude of a change in power without the direction i.e. (10W) and not (-10W, 10W). The research work considers two approaches for generating the models. The first approach examines the feature vector for subsequent changes without any gaps in between them and places these co-occurring changes into tuples which serve as representations of the charging models, as shown in (2). The second approach considers that changes that span over a longer period of time, and looks for patterns that match subsequent changes in power separated by a single gap of inactivity. These models are defined using (3).

$$model_{ng_i} = (state_{ng_{i,a}}, state_{ng_{i,b}}) \quad (2)$$

$$model_{sg_i} = (state_{sg_{i,a}}, state_{sg_{i,b}}) \quad (3)$$

where $model_{ng_i}$ is a model defined using subsequent active power changes without any gaps and $model_{sg_i}$ is a model defined using subsequent power changes separated by a single interval of inactivity.

3) Measuring Similarity

The models capture all the different state changes within a given slice i.e. time period that could be attributed to the form of significant charging activity. These representations are used to compare the similarity in power consumption between two or more time periods, and this comparison makes use of binary similarity as described by the OTUs in Table I.

As binary similarity requires the objects for comparison to be in binary feature vector form, each set of models discovered within each time period is converted into said format. The

conversion is a two-stage process with the first stage involving the extraction of all the states from each model in a given time period and placing them into a single vector.

Due to the possibility of a number of states being very similar with only a small variation in power between them it is necessary to perform a transformation of the power states in order to only capture the truly unique states. The proposed model makes use of a simple approach which rounds the power state values to the nearest tenth. For better power usage (i.e. the quality of charging stations) within a time period, both the set of no gap models and set of single gap models undergo the conversion process and a union operation is performed on the resulting sets of unique states. The two-stage process is represented by (4).

$$unique_i = states_{ng_i} \cup states_{sg_i} \quad (4)$$

where $unique_i$ is a vector of unique states for a single time period resulting from the union of the transformed power states of the no gap models $states_{ng_i}$ and the single gap models $states_{sg_i}$.

Encoding the power states as binary feature vectors requires defining a vector $unique_{ij}$ which is a union of all the unique states within both time periods, as shown in (5). The feature vector $unique_i$ for each time window is then one-hot encoded, where the presence of a power state in both $unique_i$ and $unique_{ij}$ is marked with a one and the absence is marked with a zero. The resulting representations are binary feature vectors $unique_{bi}$ which are then used in the binary similarity measures.

The *positive matches* OTU expression is used to measure and evaluate the similarity between the time periods resulting in the similarity value $similarity_{ij}$, as shown in (6). The remaining OTU expressions are used to discover which models are unique to each time window, and which models are present in both of them.

$$unique_{ij} = unique_i \cup unique_j \quad (5)$$

$$similarity_{ij} = unique_{bi} \odot unique_{bj} \quad (6)$$

Algorithm 1 The Measurement of Bag-of-Power States

Procedure Bag_of_Power_States($P_{sm}, similarity_{mn}$)

Require: Set of sequence of power states sequences, P_{sm}, P_{sn} ;

Ensure: Set of $similarity_{mn}$;

```

1   $s \in \{1, 2, \dots, S\}, m, n \in \{1, 2, \dots, M\}$ 
2  WHILE compare  $P_{sm}$  and  $P_{sn}$  DO
3       $unique_m = P_{Sng_m} \cup P_{Ssg_m}$ 
4       $unique_n = P_{Sng_n} \cup P_{Ssg_n}$ 
5       $unique_{mn} = unique_m \cup unique_n$ 
6  END WHILE
7  FOR all  $unique_{mn}$  DO
8      IF  $P_{sm} \exists (unique_m \cup unique_{mn})$ 
9           $unique_{bm} = 1$ 
10     ELSE
11          $unique_{bm} = 0$ 
12     END IF
13 END FOR
14  $similarity_{mn} = unique_{bm} \odot unique_{bn}$ 
```

where $unique_{ij}$ is the set of unique power states for both time periods i and j , and $similarity_{ij}$ is the dot product of the binary representations of power usage in time periods i and j .

A summary of the overall measurement can be concluded within the description of Algorithm 1.

V. EXPERIMENTS AND RESULTS

A synthetic dataset containing varying levels of charging load profiles was created for the feasibility and performance of the Bag-of-Power-States model. The dataset was segmented into 5 sets of load profiles, with each subset representing a different charging station, as shown in Fig. 3-7.

Station 1 is shown in Fig. 3 to have a few similarities with the other stations. The power usage in the early morning hours (12am to 6am) has some commonality with that of stations 2 and 3. Station 1 also has a few low similarities with stations 2 to 4 between 7am and 11am, and 7pm and 11pm. This station also has some commonalities with all the stations during the hour of 6pm. Lastly, there are no common power consumption patterns between 2pm and 5pm.

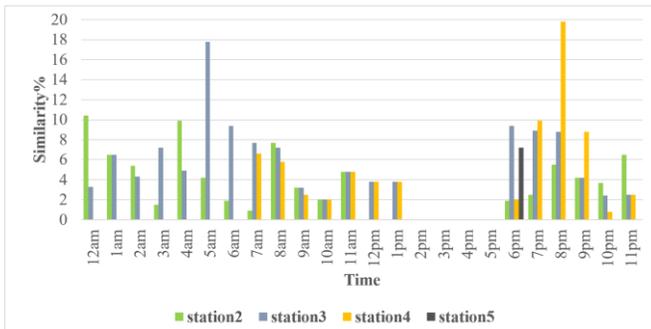


Fig. 3. Hourly Similarity Measurement of Station 1 against the Other Stations

Station 2 (Fig. 4) has similarities with stations 1 and 3 during the early morning hours, and notable similarities with station 3 for hours 1am, 2am, 10am, and 11am. Station 2 has a few similarities with the other stations except station 5 between 7am and 11am similar to station 1, with a slightly shorter evening period (8pm to 11pm). Station 1 and station 4 also seem to have some notably similar power consumption for the hour of 11am.

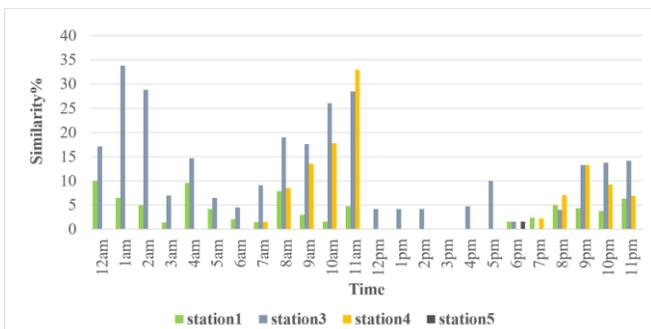


Fig. 4. Hourly Similarity Measurement of Station 2 against the Other Stations

Station 3 is shown in Fig. 5 to have a number of low similarities with station for the period 7am to 11pm. There are some notable similarities between these two stations at the hours of 10am, 11am, 2pm, 8pm, and 9pm.

Station 4 (Fig. 6) has no commonalities with the other

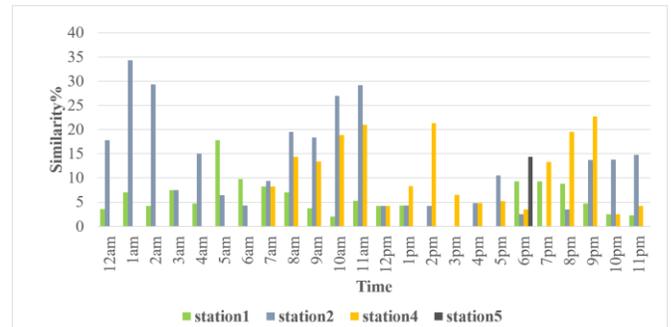


Fig. 5. Hourly Similarity Measurement of Station 3 against the Other Stations

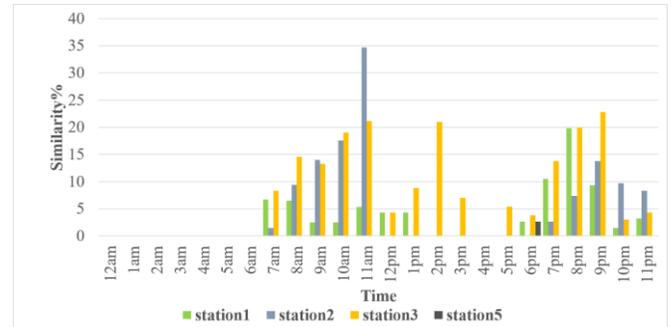


Fig. 6. Hourly Similarity Measurement of Station 4 against the Other Stations

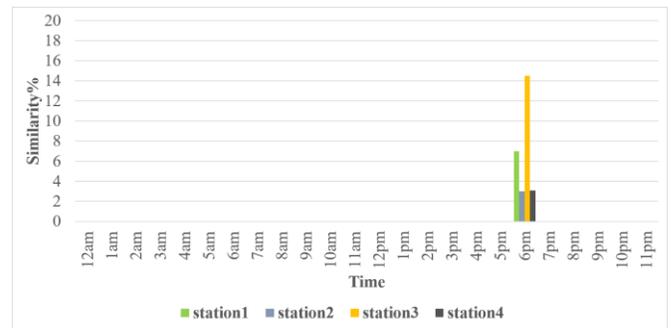


Fig. 7. Hourly Similarity Measurement of Station 1 against the Other Stations

stations for the early morning hours. However, the hours beginning from 7am have some low similarities with at least one of the other stations, except for 4pm.

Station 5 (Fig. 7) is shown to have no similarities in power usage with the other stations except for the hour of 6pm.

Each figure compares one primary substation with the other four substations providing different viewpoints of the similarities on location. In addition, the results have depicted that the proposed Bag-of-Power-States model can be put to use in order to quantify similarities between two or amongst several charging stations. The tracking of how the similarities and dissimilarities evolve over time can help to identify patterns in power consumption across various locations in order to better coordinate energy resources for ITS. The use of the slices which denote different time periods also indicate that similarities at such a resolution can help to identify key periods in time when resources must be made available to support operations.

VI. CONCLUSION

Intelligent transportation systems (ITS) are key to ensuring safety and efficient within the automotive industry. While great

progress has been made to date, there is still a wide variety of challenges to be overcome. PHEVs are a promising solution for ensuring sustainability of ITS with additional benefits for electric power and energy systems. The inclusion of the PHEVs is problematic, with charging/discharging posing threats to the stability of power grids.

This paper explored the use of power consumption data and binary vectorization in order to optimize the scheduling of power supply to charging stations. The results of experimentation using synthetic data for 5 charging stations shows the process is computationally feasible with results providing valuable insight into the dynamic nature of EV charging. Nonetheless, more efforts can be devoted to comparison of different charging datasets in the future. In parallel, practical charging data of ITS stations needs to be considered/collected in the future work, so as to better evaluate the representations of the proposed model.

ACKNOWLEDGMENT

This work has received funding from Science and Technology Project of SGCC “Research on Feature Recognition and Prediction of Typical Ice and Wind Disaster for Transmission Lines Based on Small Sample Machine Learning Method”, National Natural Science Foundation of China (No. 41911530242, 41975142), 5150 Spring Specialists (05492018012, 05762018039), Major Program of the National Social Science Fund of China (Grant No.17ZDA092), 333 High-Level Talent Cultivation Project of Jiangsu Province (BRA2018332), Royal Society of Edinburgh, UK and China Natural Science Foundation Council (RSE Reference: 62967_Liu_2018_2) under their Joint International Projects funding scheme and basic Research Programs (Natural Science Foundation) of Jiangsu Province (BK20191398).

REFERENCES

- [1] J. Kang, Z. Xiong, D. Niyato, D. Ye, D. I. Kim, J. Zhao, “Toward Secure Blockchain-Enabled Internet of Vehicles: Optimizing Consensus Management Using Reputation and Contract Theory”, *IEEE Transactions on Vehicular Technology*, 68(3): 2906-2920, 2019.
- [2] F. Malandrino, C. Casetti, C.-F. Chiasserini, “A Holistic View of ITS-Enhanced Charging Markets”, *IEEE Transactions on Intelligent Transportation Systems*, 16(4): 1736-1745, 2015.
- [3] J. E. Siegel, D. C. Erb, S. E. Sarma, “A Survey of the Connected Vehicle Landscape—Architectures, Enabling Technologies, Applications, and Development Areas”, *IEEE Transactions on Intelligent Transportation Systems*, 19(8): 2391-2406, 2018.
- [4] L. Zhu, F. R. Yu, Y. Wang, B. Ning, T. Tang, “Big Data Analytics in Intelligent Transportation Systems: A Survey”, *IEEE Transactions on Intelligent Transportation Systems*, 20(1): 383-398, 2019.
- [5] S. Kuutti, R. Bowden, Y. Jin, P. Barber, S. Fallah, “A Survey of Deep Learning Applications to Autonomous Vehicle Control”, *IEEE Transactions on Intelligent Transportation Systems*, Early Access, pp. 1-22, 07 Jan 2020. DOI: 10.1109/TITS.2019.2962338.
- [6] W. Wu, R. Li, G. Xie, J. An, Y. Bai, J. Zhou, K. Li, “A Survey of Intrusion Detection for In-Vehicle Networks”, *IEEE Transactions on Intelligent Transportation Systems*, 21(3): 919-933, 2020.
- [7] Z. Lu, G. Qu, Z. Liu, “A Survey on Recent Advances in Vehicular Network Security, Trust, and Privacy”, *IEEE Transactions on Intelligent Transportation Systems*, 20(2): 760-776, 2019.
- [8] M. G. Flammini, G. Pretticco, A. Julea, G. Fulli, A. Mazza, G. Chicco, “Statistical Characterisation of the Real Transaction Data Gathered from Electric Vehicle Charging Stations”, *Electric Power Systems Research*, 166: 136-150, 2019.
- [9] E. Xydias, C. Marmaras, L. M. Cipcigan, N. Jenkins, S. Carroll, M. Barker, “A Data-driven Approach for Characterising the Charging Demand of Electric Vehicles: A UK Case Study”, *Applied Energy*, 162: 763-771, 2016.
- [10] Y. B. Khoo, C.-H. Wang, P. Paevere, A. Higgins, “Statistical Modeling of Electric Vehicle Electricity Consumption in the Victorian EV Trial, Australia” *Transportation Research Part D: Transport and Environment*, 32: 263-277, 2014.
- [11] S. Habib, M. M. Khan, F. Abbas, L. Sang, M. U. Shahid, H. Tang, “A Comprehensive Study of Implemented International Standards, Technical Challenges, Impacts and Prospects for Electric Vehicles”, *IEEE Access*, 6: 13866-13890, 2018.
- [12] G. Bedi, G. K. Venayagamoorthy, R. Singh, R. R. Brooks and K.-C. Wang, “Review of Internet of Things (IoT) in Electric Power and Energy Systems”, *IEEE Internet of Things Journal*, 5(2): 847-870, 2018.
- [13] N. Zenina, A. Borisov, “Regression Analysis for Transport Trip Generation Evaluation”, *Information Technology and Management Science*, 16(1): 89-94, 2013.
- [14] E. I. Vlahogianni, M. G. Karlaftis, J. C. Golias, “Optimized and Meta-optimized Neural Networks for Short-term Traffic Flow Prediction: A Genetic Approach”, *Transportation Research Part C: Emerging Technologies*, 13(3): 211-234, 2005.
- [15] J. Abellán, G. López, J. de Oña, “Analysis of Traffic Accident Severity Using Decision Rules via Decision Trees”, *Expert Systems with Applications*, 40(15): 6047-6054, 2013.
- [16] R. P. D. Nath, H.-J. Lee, N. K. Chowdhury, J.-W. Chang, Modified K-Means Clustering for Travel Time Prediction Based on Historical Traffic Data, in *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (KES)*, pp. 511-521, Cardiff, UK, 8-10 Sept 2010.
- [17] E. S. Rigas, S. D. Ramchurn, N. Bassiliades, “Managing Electric Vehicles in the Smart Grid Using Artificial Intelligence: A Survey”, *IEEE Transactions on Intelligent Transportation Systems*, 16(4): 1619-1635, 2015.
- [18] H. Hu, B. Tang, X. Gong, W. Wei, and H. Wang, “Intelligent fault diagnosis of the high-speed train with big data based on deep neural networks,” *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 2106-2116, Aug. 2017.
- [19] H. Hu, B. Tang, X. Gong, W. Wei, H. Wang, “Intelligent Fault Diagnosis of the High-Speed Train with Big Data Based on Deep Neural Networks”, *IEEE Transactions on Industrial Informatics*, 13(4): 2106-2116, 2017.
- [20] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, “Long Short-Term Memory Neural Network for Traffic Speed Prediction Using Remote Microwave Sensor Data”, *Transportation Research Part C: Emerging Technologies*, 54: 187-197, 2015.
- [21] P. Wang, C.-Y. Chan, A. de L. Fortelle, A Reinforcement Learning Based Approach for Automated Lane Change Maneuvers, in *IEEE Intelligent Vehicles Symposium (IV)*, pp. 1379-1384, Changshu, China, 26-30 Jun 2018.
- [22] Z. Huang, X. Xu, H. He, J. Tan, Z. Sun, “Parameterized Batch Reinforcement Learning for Longitudinal Control of Autonomous Land Vehicles” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(4): 730-741, 2019.
- [23] B. Ko, S. Ryu, B. B. Park, S. H. Son, “Speed Harmonisation and Merge Control Using Connected Automated Vehicles on a Highway Lane Closure: A Reinforcement Learning Approach”, *IET Intelligent Transport Systems*, 14(8): 947-957, 2020.
- [24] M. Wei, M. Long, J. Yang, Research on Energy Internet Collaborative Scheduling Considering the Complex Interaction of Power System and Transportation System, in *2nd IEEE Conference on Energy Internet and Energy System Integration (EI2)*, pp. 1-4, Beijing China, 20-22 Oct 2018.
- [25] M. Rezkallah, A. Chandra, B. Singh, S. Singh, “Microgrid: Configurations, Control and Applications”, *IEEE Transactions on Smart Grid*, 10(2): 1290-1302, 2019.
- [26] M. Nasir, H. A. Khan, A. Hussain, L. Mateen, N. A. Zaffar, “Solar PV-Based Scalable DC Microgrid for Rural Electrification in Developing Regions”, *IEEE Transactions on Sustainable Energy*, 9(1): 390-399, 2018.
- [27] Q. Wang, X. Liu, J. Du, F. Kong, “Smart Charging for Electric Vehicles: A Survey from the Algorithmic Perspective”, *IEEE Communications Surveys & Tutorials*, 18(2): 1500-1517, 2016.

- [28] S. Alshahrani, M. Khalid, M. Almuhami, "Electric Vehicles Beyond Energy Storage and Modern Power Networks: Challenges and Applications", IEEE Access, 7: 99031-99064, 2019.
- [29] F. Wu, R. Sioshansi, "A Two-stage Stochastic Optimization Model for Scheduling Electric Vehicle Charging Loads to Relieve Distribution-system Constraints", Transportation Research Part B: Methodological, 102: 55-82, 2017.
- [30] Z. Liu, Q. Wu, K. Ma, M. Shahidehpour, Y. Xue, S. Huang, "Two-Stage Optimal Scheduling of Electric Vehicle Charging Based on Transactive Control", IEEE Transactions on Smart Grid, 10(3): 2948-2958, 2019.
- [31] P. Aliasghari, B. Mohammadi-Ivatloo, M. Alipour, M. Abapour, K. Zare, "Optimal Scheduling of Plug-in Electric Vehicles and Renewable Micro-grid in Energy and Reserve Markets Considering Demand Response Program", Journal of Cleaner Production, 186: 293-303, 2018.
- [32] E. B. Iversen, J. M. Morales, H. Madsen, "Optimal Charging of an Electric Vehicle Using a Markov Decision Process", Applied Energy, 123: 1-12, 2014.
- [33] E. Xydias, C. Marmaras, L. M. Cipcigan, "A Multi-Agent Based Scheduling Algorithm for Adaptive Electric Vehicles Charging", Applied Energy, 177: 354-365, 2016.
- [34] A. Verma, A. Asadi, K. Yang, S. Tyagi, "A Data-driven Approach to Identify Households with Plug-in Electrical Vehicles (PEVs)", Applied Energy, 160: 71-79, 2015.
- [35] Y. Gao, Y. Sun, X. Wang, F. Chen, A. Ehsan, H. Li, H. Li, "Multi-Objective Optimized Aggregation of Demand Side Resources Based on a Self-organizing Map Clustering Algorithm Considering a Multi-Scenario Technique", Energies, 10(12): 2144, 2017. DOI: 10.3390/en10122144.
- [36] M. B. Arias, M. Kim, S. Bae, "Prediction of Electric Vehicle Charging Power Demand in Realistic Urban Traffic Networks", Applied Energy, 195: 738-753, 2017.
- [37] A. B. Mosavi, A. Amiri, H. Hosseini, "A Learning Framework for Size and Type Independent Transient Stability Prediction of Power System Using Twin Convolutional Support Vector Machine", IEEE Access, 6: 69937-69947, 2018.
- [38] M. B. Arias, S. Bae, "Electric Vehicle Charging Demand Forecasting Model Based on Big Data Technologies", Applied Energy, 183: 327-339, 2016.
- [39] K. L. López, C. Gagné, M. Gardner, "Demand-Side Management Using Deep Learning for Smart Charging of Electric Vehicles", IEEE Transactions on Smart Grid, 10(3): 2683-2691, 2019.
- [40] Z. Chen, Z. Zhang, J. Zhao, B. Wu, X. Huang, "An Analysis of the Charging Characteristics of Electric Vehicles Based on Measured Data and Its Application," IEEE Access, 6: 24475-24487, 2018.
- [41] A. Verma, A. Asadi, K. Yang, A. Maitra, H. Asgeirsson, "Analyzing Household Charging Patterns of Plug-in Electric Vehicles (PEVs): A Data Mining Approach", Computers & Industrial Engineering, 128: 964-973, Feb. 2019.
- [42] M. Majidpour, C. Qiu, P. Chu, H. R. Pota, R. Gadh, "Forecasting the EV Charging Load Based on Customer Profile or Station Measurement?", Applied Energy, 163: 134-141, 2016.
- [43] A. S. Al-Ogaili, T. J. T. Hashim, N. A. Rahmat, A. K. Ramasamy, M. B. Marsadek, M. Faisal, M. A. Hannan, "Review on Scheduling, Clustering, and Forecasting Strategies for Controlling Electric Vehicle Charging: Challenges and Recommendations", IEEE Access, 7: 128353-128371, 2019.
- [44] M. Paradowski, "On the Order Equivalence Relation of Binary Association Measures", International Journal of Applied Mathematics and Computer Science, 25(3): 645-657, 2015.
- [45] Operational Taxonomic Unit, Wikipedia, Online Available from: https://en.wikipedia.org/wiki/Operational_taxonomic_unit, Accessed on 12 Aug 2020.
- [46] Q. Liu, K. M. Kamoto, X. Liu, M. Sun, N. Linge, "Low-Complexity Non-Intrusive Load Monitoring Using Unsupervised Learning and Generalized Appliance Models", IEEE Transactions on Consumer Electronics, 65(1): 28-37, 2019.



Qi Liu (M'11, SM'18) received the B.S. degree in Computer Science and Technology from Zhuzhou Institute of Technology, China in 2003, and M.S. and Ph.D. in Data Telecommunications and Networks from the University of Salford, UK in 2006 and 2010. His research interests include context awareness, data communication in MANET and WSN, and smart grid. His recent research work focuses on intelligent agriculture and meteorological observation systems based on WSN.



Kondwani Michael Kamoto received the B.S. in Information Technology (Software Development) from Nelson Mandela Metropolitan University, Port Elizabeth, South Africa in 2011, and M.S. degree in Computer Science and Technology in 2017, and PhD in Meteorological Information Technology in 2020, both from Nanjing University of Information Science & Technology, Nanjing, China. His research interests include IoT in electric power and power systems, and smart meter data analytics.



Xiaodong Liu (M'00, SM'17) received his PhD in Computer Science from De Montfort University and joined Napier in 1999. He is a Reader and is currently leading the Software Systems research group in the IID, Edinburgh Napier University. He was the director of Centre for Information & Software Systems. He is an active researcher in software engineering with internationally excellent reputation and leading expertise in context-aware adaptive services, service evolution, mobile clouds, pervasive computing, software reuse, and green software engineering.



Yonghong Zhang is with the School of Automation, Nanjing University of Information Science and Technology. He mainly engaged in Precision Instruments and Machinery, Pattern Recognition and Intelligent Systems, Remote Sensing Information Processing, as well as Higher Education Teaching Management and Research Work. He has led multiple major research projects of the National Natural Science Foundation of China (NSFC).



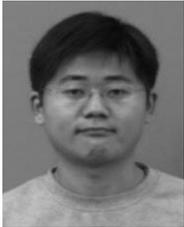
Zaiqiang Yang is with the Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science & Technology. His major research interests include Meteorological Monitoring and Disaster Warning, Intelligent Agriculture, Smart Green House, etc.



Mohammad R. Khosravi is with the Department of Computer Engineering, Persian Gulf University. He is in the field of Communication Technology, as in Electrical and Computer Engineering. He is currently an editorial board member for *Frontiers in Neuroscience*.



Yanwei Xu currently works at the Institute of Systems Engineering, Tianjin University. Their current project is 'service ecosystem, service recommendation'. In terms of ecosystem services, the concept of ecosystem extends from the biological "natural ecosystem". Borrowing the concept of ecosystem, it can be used to understand how value is created by interaction. In this research direction, the commonly used methods and techniques are probability analysis, case trajectory analysis etc. We analyze the traditional service ecosystem and online service ecosystem from the perspective of transboundary impact of service ecosystem, study the evolution law of the two ecosystems, construct the flow model of the system, and find out the law.



Lianyong Qi received the Ph.D. degree from the Department of Computer Science and Technology, Nanjing University, Nanjing, China, in 2011. In 2010, he visited the Department of Information and Communication Technology, Swinburne University of Technology, Melbourne, VIC, Australia. He is currently a Professor with the School of Information Science and Engineering, Qufu Normal University, Rizhao, China. He has chaired two NSFC projects and has published over 40 research papers in international journals, such as the *IEEE Transactions on Cloud Computing*, the *IEEE Transactions on Big Data*, the *IEEE Transactions on Computational Social Systems*, and the *IEEE Journal on Selected Areas in Communications* and international conferences, such as ICWS, ICSOC, HPCC, and TRUSTCOM. His research interests include big data and services computing.