Proceedings of the 5th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/22/20231158

The analysis of electric vehicle integration based on big data technology

Xinyao Chen

New York University, New York, NY, US 10003

Xc2017@nyu.edu

Abstract. The potential of electric vehicles (EVs) to reduce greenhouse gas emissions in the transportation sector has given the adoption of EVs a considerable boost in recent years. Concurrently, the field of big data analytics has witnessed exponential growth, providing unprecedented opportunities for extracting valuable insights and optimizing various industrial sectors. This paper presents a comprehensive overview of the intersection between electric vehicles and big data analysis. Various EV-related data sources are explored along with the discussion of data computing platforms. Followed by this, this paper analyzes different use cases of big data analysis in EVs, covering key areas such as energy management, charging infrastructure optimization, and vehicle condition monitoring, which demonstrates how big data can be crucial for the successful integration of EVs into green smart cities. Finally, the author provides future research insight and opportunities for the use of big data techniques in EV adoptions. In particular, this paper serves as a roadmap for future research in the area of data analytics in EV integration.

Keywords: big data, electric vehicles, data analysis, EV integration, energy.

1. Introduction

In recent years, electric vehicles (EVs) have emerged as a promising solution to address conventional transportation's environmental and sustainability challenges. The widespread deployment of EVs is gaining steam as the globe works to cut greenhouse gas emissions and combat climate change. At the same time, there are unheard-of chances to improve electric vehicle technology's performance, efficacy, and overall effectiveness thanks to the growing integration of big data analytic techniques.

Electric vehicles produce a vast amount of data throughout their lifespan, including areas such as vehicle performance, charging patterns, energy consumption, battery health, and driver behaviour. These diverse data streams hold immense potential to transform the EV ecosystem, optimize operations, and inform decision-making processes. By harnessing the power of big data analytics, researchers, policymakers, and industry stakeholders can unlock valuable insights to address challenges related to range anxiety, charging infrastructure optimization, battery degradation, and the overall sustainability of electric transportation systems.

The objective of this research is to explore the intersection of electric vehicles and big data, shedding light on the innovative approaches, methodologies, and applications that leverage data-driven insights to propel the adoption and efficiency of electric vehicles. Through a detailed review of the most recent data analytics technologies for integrating electric vehicles, including the cloud computing platforms and the edge computing platforms [1] and the EV grid integration (EVGI) [2], this study discusses the source of EV-related big data and its various applications aiming to highlight the transformative possibilities of big data in shaping the future of sustainable mobility.

2. Big data of electric vehicles

2.1. Electric vehicles data source

Numerous sensors in the EVs consume and produce large amounts of data [1]. The data from vehicles, drivers, charging infrastructures, or even smart city integration and market surveys constitute the big data of EVs. The data from EVs mostly comes from the electronic control unit (ECU) and battery management system (BMS) [2]. The BMS monitors various battery parameters including voltage, temperature, state of charge (SOC) and state of health (SOH). These data can be used to analyze charging patterns, optimize charging infrastructure placement, and assess the demand for charging services. The controller area network (CAN)-bus collects internal information from the vehicle such as engine speed and torque that tracks the EV status and provides insight into EV operations. In addition to the onboard data generated by the vehicles, drivers along with their smart devices also contribute to a significant portion of EVs' big data. The distance travelled, trip purpose and the arrival and departure time of drivers could be recorded or shared by the drivers. Furthermore, advanced systems can keep track of information like how much air conditioning is utilized or how a driver slows or accelerates [2]. These data could act as crucial indicators in characterizing driving behaviours. Besides, surveys and market research studies collect data on consumer preferences and attitudes related to EVs, which can help understand the factors influencing EV adoption and identify the barriers.

2.2. Smart grid and electric vehicle integration

2.2.1. Overview of SG-EV. Integration The enormous amount of data from EVs would simply remain at the concept stage without assistance in computing and analysis, and it would be challenging to improve technology. EVs are essential components of Transport Oriented Smart Cities (TOSC) since they can help to reduce emissions and the carbon footprint of cities [1]. The Smart Grid (SG), known as the next-generation power grid, builds a widely dispersed automated energy network using information, two-way, cyber-secure communication technology, and computer intelligence [3]. The EV-SG integration makes it possible for the SG to regulate EVs and holds immense potential for revolutionizing the energy and transportation sectors by extracting the necessary information to increase the efficiency and dependability of the grid. A powerful computing platform is required to support and manage the communication between EVs and SG with plenty of data generated.

2.2.2. Cloud computing. A cloud computing platform refers to a virtualized environment that provides various computing resources and services over the Internet. In cloud computing, data processing and storage primarily occur in centralized data centres located at remote locations. It offers virtually unlimited scalability as cloud providers can dynamically allocate computing resources based on demand. The cloud computing paradigm provides the necessary infrastructure, data management capabilities, and real-time communication to support EV-SG integration. However, traditional cloud computing also has major limitations:

Latency: Cloud computing typically involves data transmission over the network to remote data centres, which can introduce latency.

Bandwidth Usage: Cloud computing relies on transferring data to and from the cloud data centre, which requires significant network bandwidth, especially for large-scale data processing or storage operations.

Data Privacy and Security: Cloud computing involves transmitting data over the network to remote data centres, raising concerns about data privacy and security.

Communication: Only internet protocols (IP) could be used for communication between the cloud servers. As devices in the transportation domain utilize diverse industrial protocols and standards that differed from IP, retranslation into IP for such protocols is required for the successful transportation of data from billions of edges to data centres for storage and analysis [4].

2.2.3. Edge computing. To address the aforementioned issues with cloud computing platforms, Edge Computing could be an appropriate alternative. The idea behind edge computing is to shift a sizable percentage of computational work to the edge of the network to take advantage of edge nodes' latent computational power [1]. It contains characteristics similar to cloud computing but also includes extras like location awareness, edge deployment, and a lot of geographically dispersed nodes [5]. Edge computing reduces latency by processing data locally, near the data source. This enables faster response times and allows for real-time processing. In addition, edge computing could not only reduce bandwidth usage but also reduce the risk of data breaches as it transfers and store data locally.

3. Big data analytics for EV application

3.1. Battery management system

The Electric Vehicle Battery Management System (BMS) plays a crucial role in the performance, efficiency, and safety of electric vehicle (EV) batteries. It is responsible for monitoring, controlling, and optimizing the battery's operation throughout its lifecycle. The BMS collects and analyzes data from various sensors and modules within the battery pack, including voltage, current, temperature, and state of charge (SOC). This data enables the BMS to ensure that the battery operates within safe limits, maximizes its lifespan, and enhances overall vehicle performance. The BMS also facilitates active cell balancing, where it equalizes the charge across individual cells to maintain optimal battery health and capacity. In addition, the BMS provides critical information to the driver, such as the remaining range and charging status, through the vehicle's dashboard or infotainment system.

M. Karmawijaya et al. proposed a big data analytics architecture for sustaining the battery management system. It comprised of five architectural layers and there are defined as follows:

Data: This includes all the data sources necessary to support BMS such as battery records, temperature, current, voltage, EV battery sensors as well as information from social media and mobile Apps.

Data aggregation: In this layer, data will first be acquired from the above sources. Then, the data movement, cleaning, dividing, translating, merging, sorting, and validation will be handled by the transformation engine. Finally, databases will be used to store the managed data.

Analytics layer: It refers to data processing and analyzing procedures. Different analysis approaches and programming models should be taken based on the data's format and the analysis's goals. MapReduce is useful for tasks that can be parallelized and where the order of data processing is not critical. It allows for efficient processing of large datasets by utilizing the computing power of multiple nodes in a cluster. On the other hand, the stream computing approach allows for real-time data processing and analysis of continuous data streams. In terms of BMS, such an analysis approach could generate real-time alerts for the user when the battery is low or damaged. When it comes to battery report generation and static prediction, in-database analytics is a practical approach as it provides high-speed parallel processing, scalability, and a secure environment.

Information exploration layer: In this layer, the above analyzations are translated into visualization reports and real-time monitoring including alerts, proactive notifications, and operational performance indicators (KPIs) which display the battery's health condition and prevent unexpected circumstances.

Data governance layer: This layer emphasizes proper management of data and consideration of data security and privacy issues. To guarantee that the system complies with both BMS laws and user expectations, appropriate policies and standards should be established.

M. Karmawijaya et al. indicate that the remaining useful life (RUL) value estimation of the battery is a significant factor in verifying the performance of EV batteries for BMS. Thus, a combination of the

DOI: 10.54254/2755-2721/22/20231158

EV battery modelling method and driving pattern analysis is conducted to increase the precision of the RUL estimate. On one hand, the battery cycle-life performance and state of health ageing model was formulated with data on battery internal resistance and voltage. On the other hand, the driver's driving patterns were analyzed by transforming the driving data into a consistent vector and reorganizing it as a driving pattern. Growing hierarchical self-organizing maps (GHSOM), an unsupervised clustering approach, was used to cluster the driving patterns and their corresponding energy consumption. Together with both approaches, the RUL was calculated in the EV cloud system to provide appropriate advice for charging and route planning [6].

3.2. EV condition monitoring system

The development of electric vehicles is still in its infancy when compared to conventional cars [7]. The Electric Vehicle condition monitoring system would play a crucial part in ensuring the optimal performance, reliability, and safety of EVs. With the development of the Internet of Things (IoT), specifically the Internet of Vehicles (IoV), evaluating and monitoring the condition of cars become achievable with increasing accuracy and dependability. The EV on-board generates various data as an essential guide for the safety and stability evaluation, which includes parameters for the battery status (voltage, current, temperature and SOC, insulation resistance), the motor status (motor speed, voltage, current, temperature) and the driving parameters. In addition to the data directly collected from EVs, mobile communication technology, satellite positioning technology (RS), and geographic information system (GIS) would improve the information construction [8].

X. Feng and J. Hu established a deep network data mining model based on Hadoop to analyze vehicle behaviour. The system is divided into a perception layer, transmission layer and application layer. The perception layer refers to the vehicle terminal that can receive vehicle operational data and carry out remote control commands. In the transmission layer, data collected is transmitted to the remote monitoring centre via a 3G module in real time. Finally, the application layer oversees observing the monitoring parameters and visualizing the internal association data relationship through data mining and analysis [8].

3.3. Smart city integration

3.3.1. Electric taxi and ride-hailing. As cities strive to reduce greenhouse gas emissions and combat air pollution, the electrification of public transportation presents a crucial opportunity for a shift toward a more sustainable and environmental-friendly mobility solution. With the rapid growth of EV technology and government incentive policies worldwide, taxis and ride-hailing platforms have increasingly adopted electric vehicles as an eco-friendly substitute for traditional combustion engine vehicles. While the adoption of electric taxis brings numerous benefits, some challenges and shortcomings posed questions of its viability:

Limited Range and Charging Infrastructure: Electric taxis typically have a limited driving range compared to traditional taxis. The need for frequent recharging can disrupt operations and lead to downtime.

Longer charging time: Charging an electric taxi takes considerably more time than refuelling a conventional taxi. This can result in longer waiting times for drivers which may affect their productivity and potentially reduce the number of trips they can complete in a day.

Integration with existing infrastructure: The transition to electric taxis requires adapting existing infrastructure, including charging stations and maintenance facilities. Upgrading and integrating these elements to support electric taxis can be a complex and costly process.

C. -M. Tseng et al. formulated a Markov Decision Process (MDP) to stimulate computerized electric taxi service tactics. The optimal policy of the MDP is obtained by using a sizable dataset of New York City taxi trips. Dynamic programming embodied with MDP is utilized to maximize the total expected net revenue for taxi drivers. The parameters for MDP are driving speed network, passenger pick-up probability, passenger destination probability, energy consumption taxi net revenue and charging rate.

From an analysis of the NYC taxi travel dataset's one-day data for January 9, 2013, the authors find that a taxi driver can earn at most among the top 0.01% thanks to MDP's technique for optimizing taxi service. A similar evaluation was conducted on the data of electric taxis. The result indicated that with more battery capacity (e.g., 50 kWh, 70 kWh), the net income earned by the driver of an electric taxi can be on par with that of a regular taxi driver. Thus, the viability of electric taxi integration is verified if the battery capacity for the electric taxi is above 50 kWh [9].

3.3.2. EV charging station placement. The strategic placement of EV charging stations is crucial for facilitating the widespread adoption of EVs. A comprehensive charging infrastructure plan should be developed in coordination with stakeholders, including government entities, utility companies, businesses, and EV owners. There are many studies investigating the place of EV charging stations from the distributions of EVs [10, 11]. Y. Yuan et al. took a different approach to propose charging station placement from a ride-hailing service perspective from big data networking. The authors took their inspiration from the content distribution network (CDN). A crucial component of a content distribution network is the placement of numerous caching servers in regions of the network where users are more likely to access them. Based on the ride-hailing service data collected from DiDi company that covered 77 km^2 of Chengdu, China, the author proposed a ride-hailing charging station placement (RHCSP) approach that takes into account the road access frequency, high-power consumption and high-power demand intensity. To determine the coordinates of probable charging stations, a hotspot path mining method was first developed. The algorithm consisted of the gridding method, which quantizes the space into many grids and then clusters on these grid cells, and hotspot path mining, which map all data points into corresponding grids with unique latitude and longitude. Then, the coordinates of all sections where power consumption or power demand is higher than a specific threshold were extracted. Finally, the union of these coordinates is taken into consideration. By eliminating the points with close distance, the coordinate point set of the charging station site was obtained. Such a method was proved to have a high hit rate and low distance cost [12].

4. Future research direction

Although there are plenty of studies about big data applications on EVs, the research of big data analytics for EVs is still in its early stage. The analyzed data have not yet been used for decision-making purposes in many studies. Investigating the integration of EVs with smart city initiatives could be a significant research area that improves the current policies and enhance overall urban mobility by analyzing the role of EVs in transportation networks, energy systems, and urban planning. The possibility of taking EV as an emergency backup supply for home, often known as vehicle-to-home (V2H), or for another vehicle (V2V) might be realized in future research [2]. While big data has enormous benefits, data privacy and security issues are still a concern. There is a need for developing privacy-preserving techniques and secure data-sharing protocols to protect EV data while still enabling data-driven research and innovation. In addition, the edge computing platform mentioned in Section 2 is still in the concept phase. To realize the edge technology, various challenges need to be overcome including identifying the suitable programming models and workflow partitioning, solving the scalability problem, establishing comprehensive security and risk assessment processes, and creating an orchestration framework to allow accurate alignment and mapping to the resource pool [4].

5. Conclusion

This article provides an overview of the use of big data analysis in electric vehicle integration. The sources of EV-related big data including the vehicles, drivers, batteries, and those linked to smart cities are discussed. Cloud computing and edge computing are two possible computing platforms for these data. While the cloud computing platform could offer virtually unlimited scalability, it involves several shortcomings including latency, bandwidth usage, data privacy issues, and communication problems. The edge computing platform may address the above issues but is still in the concept phase and more research in this area need to be conducted. The development of an integrated battery management system,

DOI: 10.54254/2755-2721/22/20231158

the design of a vehicle condition monitoring system based on Internet of Things technology, a test of the viability of the popularization of electric taxis based on the Markov Decision Process model, and EV charging stations placement for ride-hailing services using content distribution networking approach were all covered in this article's review of various applications of the big data analysis approach for EVs. Big data technology is used in EVs in more applications than only those covered in this article. It would be a good idea to look into more recent applications. Big data offers significant benefits for the society and the economy. Research in the future might not only concentrate on the data analysis side for EV features but also examine methods to realize analyzed data into decision-making objectives and enhance data privacy and security.

References

[1] Hussain, M. M., Beg, M. M. S., Alam, M. S., Krishnamurthy, M. & Ali, Q. M., Computing Platforms for Big Data Analytics in Electric Vehicle Infrastructures [C] 2018 4th International Conference on Big Data Computing and Communications (BIGCOM), Chicago, IL, USA, 2018, pp. 138-143.

[2] Kisacikoglu, B. Li, M. C., Singh C. Liu, N. & Erol-Kantarci M., Big Data Analytics for Electric Vehicle Integration in Green Smart Cities [J] IEEE Communications Magazine, vol. 55, no. 11, 2017, pp. 19-25.

[3] Fang, X., Misra, S., Xue G. & Yang, D., Smart Grid — The New and Improved Power Grid: A Survey [J] IEEE Communications Surveys & Tutorials, vol. 14, no. 4, 2012, pp. 944-980.

[4] Hussain, M. M., Beg, M. S., Alam, M. S., & Laskar, S. H., Big Data Analytics Platforms for Electric Vehicle Integration in Transport Oriented Smart Cities: Computing Platforms for Electric Vehicle Integration in Smart Cities.[J] International Journal of Digital Crime and Forensics (IJDCF), vol. 11, no.3, 2019, pp.23-42.

[5] Osanaiye, O., Chen, S., Yan, Z., Lu, R., Choo, K. -K. R. & Dlodlo, M., From Cloud to Fog Computing: A Review and a Conceptual Live VM Migration Framework [J] IEEE Access, vol. 5, 2017, pp. 8284-8300.

[6] Karmawijaya, M. I., Nashirul Haq, I., Leksono, E. & Widyotriatmo, A., Development of Big Data Analytics Platform for Electric Vehicle Battery Management System [C] 2019 6th International Conference on Electric Vehicular Technology (ICEVT), Bali, Indonesia, 2019, pp. 151-155.

[7] Wang, Z. & Li, G., 3D Modeling Technology of Mobile Big Data Application in the Electric Vehicle Condition Monitoring System [J] Wireless Communications and Mobile Computing, vol. 2022, 2022, pp. 13.

[8] Feng, X. & Hu, J., Research on the identification and management of vehicle behaviour based on Internet of things technology [J] Computer Communications, vol. 156, 2020, pp. 68-76.

[9] Tseng, C. -M., Chau, S. C. -K. & Liu, X., Improving Viability of Electric Taxis by Taxi Service Strategy Optimization: A Big Data Study of New York City [J] IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 3, 2019, pp. 817-829.

[10] Deb, S., Tammi, K., Kalita, K., Mahanta, P. Impact of Electric Vehicle Charging Station Load on Distribution Network [J] Energies, vol. 11, no.1, 2018, pp.178.

[11] Hosseini, S., Sarder, MD. Development of a Bayesian network model for optimal site selection of electric vehicle charging station [J] International Journal of Electrical Power & Energy Systems, vol. 105, 2019, pp. 110-122.

[12] Yuan, Y., Zhu, K., Zhang, L. & Ge, Y., An EV Charging Station Placement for Ride-hailing Service: from Big Data Networking Perspective [C] 2020 International Conference on Computing, Networking and Communications (ICNC), Big Island, HI, USA, 2020, pp. 235-239.