International Journal of Global Economics and Management
ISSN: 3005-9690 (Print), ISSN: 3005-8090 (Online) | Volume 6, Number 2, Year 2025
DOI: https://doi.org/10.62051/ijgem.v6n2.02
Journal homepage: https://ijgem.org

Marketing Path Analysis Based on Spatio-Temporal Big Data of Internet of Vehicles

Xuehong Wang [1, 2], Xinyu Ma [1, 2], Fan Zhang [1, 2, *]

[1] China Automotive Technology and Research Center Co., Ltd. Tianjin, China
[2] China Auto Information Technology (Tianjin) Co., Ltd. Tianjin, China
*Corresponding Author: m13502037247_3@163.com

ABSTRACT

With the rapid development and popularization of vehicle networking technology, the spatio-temporal big data generated by vehicles has brought new opportunities to the field of marketing. This paper aims to explore the marketing path analysis based on the spatiotemporal big data of the Internet of vehicles, and provide accurate marketing strategies for enterprises by digging deeply into vehicle driving data, driving habits and location information, etc. First, this paper introduces the basic concepts and characteristics of spatiotemporal big data of the Internet of vehicles, then introduces the collection and processing of the data of the Internet of vehicles, and then builds the feature engineering and the text based on CatBoost model to predict the future of consumers. Finally, this paper analyzes the application strategy path of spatiotemporal big data in marketing, including precision marketing, personalized recommendation and precision advertising based on user portraits. The method proposed in this paper can not only improve user experience, but also create higher business value for enterprises.

KEYWORDS

Internet of Vehicles; Spatiotemporal big data; Marketing path; Personalized recommendation; Targeted advertising

1. INTRODUCTION

As a typical application of Internet of Things technology in the field of transportation, vehicle networking realizes multi-dimensional information interaction and collaborative sharing between vehicle and vehicle (V2V), vehicle and infrastructure (V2I) and vehicle and person (V2P) through the deep integration of vehicle and Internet and communication network. With the rapid development of vehicle networking technology, the scale of spatio-temporal big data generated by vehicles is growing exponentially. The data dimension not only includes dynamic information such as vehicle's driving track, speed, acceleration and residence time, but also covers the user's driving behavior pattern, consumption preference and other deep-seated characteristics. This trend has brought unprecedented opportunities and challenges to the field of marketing. Specifically, the generation and utilization of massive spatio-temporal data provides a multi-source heterogeneous data basis for automobile marketing [1], making it possible to analyze consumption scenario preferences, optimize marketing strategies and predict user behavior [2], thus providing theoretical basis and technical support for enterprises to achieve precision marketing and personalized services [3].

The spatiotemporal big data analysis of the Internet of vehicles has become a research hotspot of common concern in academia and industry at present. The Internet of Vehicles (V2X) technology integrates on-board sensors, global positioning system (GPS) and multi-mode communication modules to achieve real-time acquisition and transmission of vehicle location information, driving trajectory, driving behavior and other data. According to Statista statistics show that in 2023, the global vehicle networking market has exceeded 160 billion US dollars, and the average daily data generation exceeds 10TB. These massive data contain rich user space-time behavior characteristics, providing new methodological support for urban traffic congestion analysis, traffic management optimization and other research. For example, by introducing the unsupervised clustering algorithm K-Shape, the literature [4] constructs a series analysis model of deviation values between the predicted trajectory and the real trajectory, and takes the velocity variation coefficient as the key characteristic index, which significantly improves the accuracy of traffic event prediction. The experimental results show that the proposed method has high accuracy and robustness in event state recognition. Based on vehicle track data and vector road network data, the paper [5] proposes a spatial correlation and propagation mining algorithm for road section congestion in large-scale urban road network, and realizes accurate positioning and quantitative analysis of high-correlation road section congestion. The research proves that the algorithm can still effectively depict the vehicle path in complex scenarios such as multi-line parallel sections or track point drift, and provides reliable technical support for the traffic congestion situation monitoring of large-scale road networks. In order to solve the problems of heterogeneity of spatio-temporal data of vehicle-networking in urban scenes and low communication efficiency within the coverage of a single infrastructure, reference [6] proposed a spatiotemporal data analysis and accessibility optimization method for vehicle-networking. Simulation results show that the proposed method can effectively predict the connectivity intensity between vehicles, significantly reduce communication redundancy and reduce the load pressure on roadside infrastructure. These research results not only validate the application value of spatiotemporal big data in the field of traffic management, but also provide an important theoretical reference and practical basis for the subsequent marketing analysis of automotive products.

Based on the above factors, this paper proposes a method based on "Marketing path analysis based on spatiotemporal big data of the Internet of Vehicles", which reveals the driving behavior pattern, consumption preference and travel characteristics of users by mining and analyzing the massive spatiotemporal data generated in the environment of the Internet of vehicles, so as to provide precise marketing strategy support for enterprises.

2. RESEARCH METHODS

This paper carries out analysis according to the five-stage framework of "definition and characteristics of spatio-temporal big data -- data collection and processing -- feature extraction -- model construction -- strategy generation". The first stage is to define and analyze spatio-temporal big data of the Internet of vehicles according to the characteristics of large amount of data, high real-time performance, diverse dimensions and high spatio-temporal correlation. The second stage is the data collection and processing of the spatio-temporal big data of the vehicle network, which mainly uses the data cleaning method to remove the duplicate, error and incomplete data. The third stage is feature construction, mainly using the logical operation of data fields to construct new data features that can be used by the post-text Catboost model. The fourth stage is model construction, mainly to build Catboost model to predict the future consumer demand. The fifth stage is strategy generation, mainly auto product sales strategy generation. The analysis roadmap of this paper is shown in Figure 1 below.

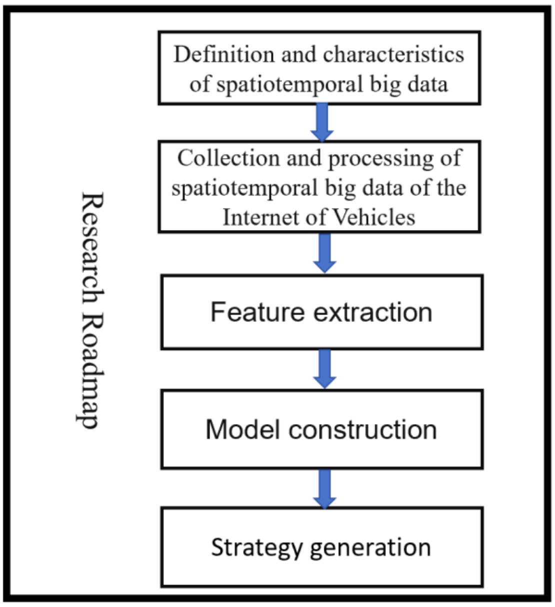

Figure 1. This paper analyzes the roadmap

2.1. Definition and Characteristics of Spatiotemporal Big Data of the Internet of Vehicles

The spatiotemporal big data of the Internet of vehicles refers to the large-scale data collection generated by vehicles, roads, and traffic facilities in the environment of the Internet of vehicles. These data all have temporal and spatial attributes. Its content covers basic information such as the vehicle's driving track, speed, acceleration, position and direction, as well as deep-seated information such as the driving habits, preferences and behavior patterns of the vehicle owners. These data have the characteristics of large amount of data, high real-time, diverse dimensions, and high spatial and temporal correlation, which can comprehensively describe the running state of vehicles and the behavioral characteristics of users.

(1) Large amount of data: car networking environment car networking system is generating a large amount of data at all times, including vehicle location information, driving speed, driving trajectory, etc. These data constitute the basis of car networking spatiotemporal big data, and the scale of data continues to expand with the increase in the number of vehicles.

(2) Strong real-time: The spatiotemporal big data of the Internet of vehicles has a strong real-time nature, which can reflect the running state of the vehicle and the behavioral changes of the user in real time, providing enterprises with timely and accurate marketing information and timely data support for the development of marketing strategies for enterprises.

(3) Diversity: The spatiotemporal big data of the Internet of vehicles comes from a wide range of sources, including vehicle sensors, GPS positioning system, communication network data, vehicle location and driving information and other data information. These data types are diverse, including numerical, text, image, etc., which can reflect the needs and characteristics of users from multiple dimensions.

(4) High correlation between time and space: each record in the spatio-temporal big data of the Internet of vehicles accurately marks the time and space information, and can accurately reflect the state and behavior of the vehicle at a specific time point and location.

2.2. Collection and Processing of Spatio-Temporal Big Data of the Internet of Vehicles

2.2.1. Data collection

The collection of spatiotemporal big data mainly relies on a variety of advanced technical means, including vehicle sensors, GPS positioning system, communication network and cloud computing platform. Vehicle sensors (such as inertial measurement unit, radar, Lidar and camera, etc.) are used to collect real-time vehicle driving status information, including speed, acceleration, steering Angle, braking state, fuel consumption or power consumption and other dynamic data. GPS positioning system accurately obtains real-time geographic location information of vehicles through satellite signals, and combines high-precision map data to further provide vehicle's driving trajectory, path planning and location context information. Communication networks (such as 4G/5G, V2X, DSRC, etc.) are responsible for efficient transmission of multi-source heterogeneous data collected by vehicle sensors and GPS to ensure real-time and integrity of data, and gather it to the cloud or edge computing platform for processing and analysis [7]. The schematic diagram of data acquisition for the Internet of vehicles is shown in Figure 2 below.

The connected vehicle system can also further enrich the dimension and scale of spatio-temporal big data by interacting with the data of road infrastructure (such as intelligent traffic lights, roadside units) and other vehicles, providing a solid data foundation for traffic management, driving behavior analysis, intelligent navigation and commercial applications.

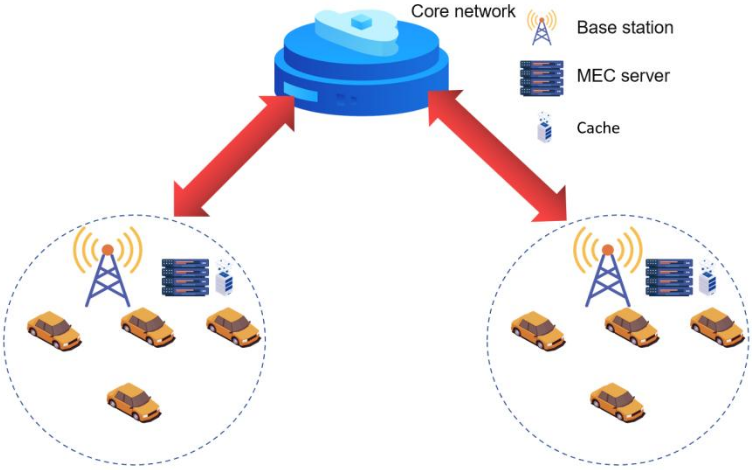

Figure 2. Schematic diagram of Internet of Vehicles data collection

2.2.2. Data processing

In the process of data collection, it is necessary to pay attention to the quality and integrity of data to ensure the accuracy and reliability of data. At the same time, due to the large scale and diverse types of big data in the space and time of the Internet of vehicles, efficient data processing technology needs to be used for cleaning, integration and analysis. Common data processing technologies include data cleaning, data fusion, data mining and so on. Data cleaning is used to remove duplicates, errors, and incomplete data. Data fusion is used to integrate data from different sources to form a unified data view. Data mining is used to mine valuable information and patterns from massive data.

2.3. Feature Extraction

Feature combination is a key technology in feature engineering, which generates new derived features by mathematical or logical operations on existing feature variables. For example, given two original features (such as a and b), new features can be constructed by operations such as addition (a+b), subtraction (a-b), multiplication (a*b), or division (a/b, taking care to deal with cases where the divisor is zero). This feature combination strategy is able to reveal potential nonlinear relationships and interaction effects in the data that may not be intuitive or easy to capture in the original feature space. By introducing feature combination, the expressiveness of the model is enhanced, enabling it to more accurately depict the inherent laws and structures of the data, thus significantly improving the performance of the model in various prediction tasks [8].

(1) Feature combination

For the features with the same attributes, the clustering method is used to gather the related attributes together, such as "year", "month", "day", "time" and so on into "date". The calculation formula for combining features is shown in formula (1).

$$s = set(\text{if i } X \text{ for i in word}) \tag{1}$$

The meaning of formula (1) is to cycle through each feature, combining features that have the same meaning into a group.

(2) Feature reduction

The goal of feature reduction is to optimize model performance and reduce computational complexity by eliminating irrelevant or redundant feature indexes. Whether to delete a feature can refer to formula (2), which can quantify the contribution of the feature to the model and provide a basis for feature screening.

$$\text{g}=(1-\frac{1}{n})\times 100\% \tag{2}$$

In formula (2), n represents the sum of different values in all eigenvalues of a feature index of the data source.

2.4. Model Construction

The internet of vehicles data usually contains a large number of category features, and traditional machine learning models need complex preprocessing (such as unique heat coding or label coding) when processing these features, while CatBoost can directly process category features without additional coding steps. The base model of CatBoost adopts a symmetric tree structure, which is unique in its ability to generate new class features by using the combination of relationships between features (i.e. extending the feature dimension through the combination of features), thus enhancing the expressiveness of the model. In addition, CatBoost introduces L2 regularized hyperparameters at the end of the model, which not only helps to improve the prediction accuracy of the model, but also effectively prevents overfitting [9].

During feature recombination, the number of recombined features increases exponentially with the number of class features, especially in large datasets. However, in practical engineering applications, it is impossible to enumerate all possible feature combinations. The CatBoost model adopts an efficient strategy: only when the current tree needs to choose a new split point, the optimal feature combination is dynamically computed in the form of greedy algorithm [10]. The specific calculation method of this greedy segmentation method is shown in formula (3).

$$Gain=\frac{1}{2}\left[\frac{G_L^2}{H_L+\lambda}+\frac{G_R^2}{H_R+\lambda}-\frac{(G_L+G_R)^2}{G_L+G_R+\lambda}\right]-\gamma \tag{3}$$

After the reorganization of the eigenvalues is completed, the Gini coefficient of each possible segmentation scheme needs to be evaluated. The segmentation scheme with the smallest Gini coefficient will be selected as the optimal segmentation point of the node, and its calculation model is shown in formula (4):

$$y_i = \varphi(x_i) = \sum_{k=1}^{k} f_k(x_i),\ fk \in F \tag{4}$$

Where k is the number of basic tree models, F is the basic tree model, on this basis, assuming that the learning errors of the basic model tree are independent of each other, add a new function F(t) to the model, then the learning goal of CatBoost model is to find $f_t(x)$ and minimize the objective function, as shown in the following formula (5):

$$\hat{y}_i^{(t)} = y_i^{(t-1)} + f_t(x_i) \tag{5}$$

The learning of the i-th sample in the t-round model needs to combine the results of the previous t-1 round to learn the current round number model, so the objective function of the i-th sample during the t-round is formula (6).

$$\begin{aligned} L^{(t)} &= \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t)}\right) + \Omega(f_t) \\ L^{(t)} &= \sum_{i=1}^{n} l\left(y_i, y^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \end{aligned} \tag{6}$$

2.5. Strategy Generation

Based on the above calculation results, enterprises can adopt the following three strategies according to the actual situation.

(1) Enterprises can build user portraits of car owners through vehicle networking data, laying a solid foundation for precision marketing. The user portrait is a comprehensive description of multidimensional information based on the basic attributes of the owner (such as age, gender, occupation), behavioral characteristics (such as driving habits, travel frequency) and interests and hobbies, which can help the enterprise to gain in-depth insight into the needs and preferences of the owner. By analyzing the vehicle's driving track, speed, location and other spatiotemporal data, combined with the owner's driving behavior pattern and consumption preferences, enterprises can accurately mine the owner's travel needs, consumption habits and interest tendencies, so as to build a comprehensive and dynamic user portrait. This portrait can not only support the formulation of personalized marketing strategies, but also provide data-driven decision-making basis for enterprises to optimize product design and service experience.

(2) Based on the spatiotemporal big data of the Internet of vehicles, enterprises can provide highly personalized travel recommendations and value-added services for car owners. Through the in-depth analysis of vehicle trajectories, driving behavior patterns and spatio-temporal data, enterprises can accurately understand the travel needs and preferences of car owners, so as to customize the best travel plan for them. For example, for car owners who often drive on congested roads, companies can recommend more efficient and smooth alternative routes in real time; When car owners need to park, the system can provide vacancy prediction and navigation services in nearby parking lots based on historical data and real-time information, greatly improving travel efficiency.

Enterprises can also use the data to provide personalized entertainment and life service recommendations for car owners. By analyzing car owners' driving time, location information and historical behavior data, companies can predict car owners' points of interest and potential needs, so as to recommend services such as music, movies and dining that fit their preferences. For example, when the car owner is about to arrive at the business district, the system can actively push the information of nearby restaurants, coffee shops and other leisure places, and provide personalized coupons or discount information based on the user's consumption habits, to further improve the user experience and satisfaction. This kind of precise service based on data can not only enhance user stickiness, but also create more business value for enterprises.

(3) Based on the spatiotemporal big data of the Internet of vehicles, enterprises can achieve accurate advertising and marketing strategy optimization. By deeply analyzing the car owner's driving trajectory, location information and activity area, the company can accurately insight the car owner's travel habits and interest range, so as to push highly relevant advertising content for them. For example, when the car owner frequently appears in a business district, the system can push the promotion activities or preferential information of the merchants in the area in real time. When the car owner approaches a tourist attraction, the enterprise can take the initiative to provide scenic spot introduction, ticket discounts and peripheral service recommendations to enhance the practicality and attractiveness of the advertisement.

The spatiotemporal big data of the Internet of vehicles also supports the targeted delivery and effect evaluation of advertising. By integrating the basic attributes of car owners (such as age, gender, occupation) and deep behavioral characteristics (such as driving habits, consumption preferences), companies can accurately target ads to target groups with specific needs and interests, significantly improving the conversion rate and delivery efficiency of ads. At the same time, enterprises can scientifically evaluate advertising effects by monitoring key indicators such as click-through rate and conversion rate of advertisements, and continuously optimize advertising strategies based on data analysis results to maximize marketing effects. This data-driven precision marketing model can not only improve the user experience, but also create higher business value for enterprises.

3. MARKETING CHALLENGES AND COUNTERMEASURES BASE ON SPATIOTEMPORAL BIG DATA OF THE INTERNET OF VEHICLES

3.1. Data Privacy and Security Challenges

In the process of using spatiotemporal big data of the Internet of vehicles for marketing, data privacy and security issues have become key challenges that enterprises must deal with. The Internet of vehicles data contains highly sensitive information such as car owner's driving trajectory and real-time location, which, if handled improperly, may lead to personal privacy disclosure and security risks. Therefore, enterprises must strictly implement data privacy and security protection measures in the whole life cycle of data collection, storage, processing and analysis. Through the use of data encryption technology, strict access control mechanism and anonymous processing means to ensure the security and confidentiality of data, so as to achieve precision marketing, while protecting the privacy rights of car owners and data security.

3.2. Data Quality and Integration Challenges

The spatiotemporal big data of the Internet of vehicles has the characteristics of a wide range of sources and diverse types, and data quality and integration are the key links that enterprises need to focus on in the application. Since the multi-source data may have problems such as non-uniform format, missing or wrong data, enterprises must carry out strict data cleaning and integration in the data collection and processing stage to ensure the accuracy and integrity of the data. In addition, enterprises also need to establish a unified data standard and specification system to provide standardized support for the integration and in-depth analysis of multi-source data, so as to improve the availability of data and analysis efficiency.

4. ANALYSIS AND DISCUSSION

This paper deeply discusses the marketing path analysis method based on the spatiotemporal big data of the Internet of vehicles. Through the systematic analysis of the basic concept, technical characteristics and application path of the spatiotemporal big data of the Internet of vehicles in marketing, this paper draws the following conclusions: The spatiotemporal big data of the Internet of vehicles has brought great potential and value to the marketing field. Enterprises can deeply explore the vehicle's driving track, speed, location and other spatiotemporal data as well as the owner's driving habits, preferences and other behavioral patterns, accurately insight into the needs and preferences of car owners, so as to provide them with highly personalized recommendations and services. In addition, the spatiotemporal big data of the Internet of vehicles also provides strong support for enterprises to achieve accurate advertising and marketing effect evaluation, which can significantly improve the accuracy, efficiency and conversion rate of advertising, and create higher commercial value for enterprises.

REFERENCES

[1] DU Mengru. Application of Digital Strategy in Automobile Marketing [J]. Automotive Test Report, 2024, (13): 158-160.

[2] PuXingJin. W M new energy car marketing strategy optimization research [D]. Guangxi university, 2024. The DOI: 10.27034 /, dc nki. Ggxiu. 2024.000611.

[3] Li Xinyan. Wei to automobile marketing strategy optimization research [D]. Yanshan university, 2024. The DOI: 10.27440 /, dc nki. Gysdu. 2024.000624.

[4] Cao Hongli. Research on Highway Traffic Event Detection and Spatiotemporal Impact Estimation Based on Vehicle Trajectory Data [D]. Southwest jiaotong university, 2023. DOI: 10.27414 /, dc nki. Gxnju. 2023.002293.

[5] ZHOU Qifan. Research on Spatial-temporal association rules mining of Urban Traffic congestion based on vehicle trajectory data [D]. Southwest jiaotong university, 2023. DOI: 10.27414 /, dc nki. Gxnju. 2023.003732.

[6] CHENG Jiujun, Yuan Guiyuan, Cui Jie, et al. Spatiotemporal data analysis and accessibility method of vehicle networking in urban scene [J]. Journal of Communications and Technology, 2021, 42 (06): 52-61.]

[7] Yi Chenyu, Tan Zhirong, Cao Peng. From data acquisition to perceive - automated driving environment perception technology application review [J]. Journal of automobile repair and maintenance, 2025, (02) : 95-99. The DOI: 10.13825 / j.carol carroll nki motorchina. 2025.02.023.

[8] ZHANG Fan, GUO Yaxin, Yang Jing, et al. Research on electric energy prediction based on GBDT+ feature engineering method [J]. Electronic Quality, 2020, (01): 1-4.

[9] LAN Yutian, Yao Wei, Zhang Wendong, et al. A Two-stage Lightweight Transient Stability Intelligent Evaluation Method for Novel Power System Based on CatBoost [J]. Shandong electric power technology, 2024, 51 (02) : 1-10. DOI: 10.20097 / j.carol carroll nki issn1007-9904.2024.02.001.

[10] Chen Hao. Research on Online Stability Evaluation Method of Power System based on Machine Learning [D]. The three gorges university, 2022. DOI: 10.27270 /, dc nki. Gsxau. 2022.000499.