DOI: 10.21917/ijsc.2025.0520

FRAUDULENT TAXI DRIVER DETECTION: A REVIEW

Zainab S. Al-Sudani[1], Musaab Riyadha[2] and Ali A. Titinchi[3]
[1,2]Department of Computer Science, Mustansiriyah University, Iraq
[3]Department of Electrical and Computer Engineering, College of Engineering, University of Nizwa, Sultanate of Oman

Abstract

This review includes the methods and tactics involved in the detection of dishonest taxi drivers, explaining their vital significance within the transport sector. It looks into different studies and strategies such as detecting outlier patterns, deployment of AI, and examining social structures in order to expose wrongful actions and conduct of rogue drivers. Constant use of big transactional databases, GPS systems and details also raises the quality of fraud prevention mechanisms. The current study assesses the roles of clustering, classification and outlier detection respectively in identifying anomalies and other related frauds within the taxi service. The incorporation of time, place and money factors has been proven to be very important as it enhances the effectiveness and speed of fraud detection systems. Change is always accompanied by problems and in this instance it is the problem to adapt to the new conditions that move sovereignty further in the process and delay the answer to fraud detection challenges. In order to achieve these goals persistent efforts on development and research will be necessary. To conclude this review, tactical methods and approaches in the detection of fraudulent taxi drivers were discussed together with how service providers and other transportation agencies can reduce cost through enhancement of passenger's security and the reputation of the sector.

Keywords:

Taxi Fraud, Global Position System (GPS), Machine Learning, Deep Learning, Decision Tree Algorithms, Shortest Path Problem

1. INTRODUCTION

. Fraudulent behavior by taxi drivers, such as taking unnecessary detours to overcharge passengers, continues to be a significant issue that undermines trust in taxi services and poses financial and safety risks to passengers [1, 2]. Several taxi companies have integrated GPS devices into their fleet of cars to monitor the behavior of their drivers. These devices record the trajectory of taxis, showing spatial coordinates (x, y) and time stamps, which enable a close analysis of routes and travel patterns [3, 4], for example. This tracking ensures that service quality improves and accountability in service provision within the industry. For instance, according to [5], fraud pattern identification is hinged on an analysis of GPS data and route deviations. Such detection requires spatial awareness and sophisticated methods to handle nuances like shortcuts or detours. The main problem in fraudulent detection is to address sufficiently how systems handle edge cases that appear fraudulent but have legitimate explanations, potentially creating unnecessary disputes and eroding driver trust. A GNN-based method that enhances fraud detection accuracy by learning from contextual relationships and selectively combining relevant information to better distinguish between legitimate and fraudulent behavior [6].

Fraudulent behavior has been overcome by several machine learning techniques over the years. Traffic flow and density can be modelled in road networks using clustering algorithms such as BIRCH or DBSCAN, and abnormal patterns that indicate fraud can be identified very easily [7]. Supervised learning methods, including decision trees and ensemble models, have lately been used more frequently to classify fraud cases versus legitimate behavior using trajectory and transactional data. Besides, unsupervised learning methods, such as density-based clustering, are suitable for anomaly detection problems, which do not require labelled datasets. This makes them very useful for practical applications when the amount of labelled data could be small. However, Fig.1 presents a taxonomy summarizing the different methods used for detecting fraudulent taxi drivers. The approaches are classified based on detection techniques (e.g., rule-based, machine learning, deep learning), data sources (e.g., GPS trajectory, fare logs, passenger feedback), and feature types (e.g., spatial, temporal, behavioral). This structure helps organize existing research and highlights key areas of this paper.

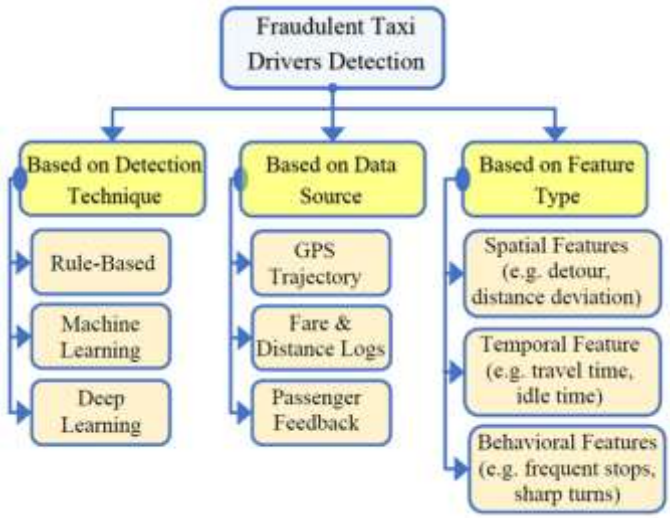

Fig.1. Taxonomy of different methods used for detecting fraudulent taxi drivers

The Fig.2 illustrates the overall architecture of a fraud taxi driver detection system. The data collection process is initiated at sensors and GPS, which gather data about the trip regarding the route and timings taken. This raw data is then cleaned and organized in the pre-processing stage, ensuring accuracy and consistency. The fraud detection module follows, employing machine learning algorithms and rule-based methods to identify suspicious behaviors. Any detected anomalies trigger the alert/notification stage, where operators are notified in real-time. The evaluation stage then assesses

the accuracy and effectiveness of these alerts, feeding back into the detection models for refinement. Finally, all data, including alerts and evaluation outcomes, is stored in the data storage stage, providing a valuable resource for future analysis and system optimization.

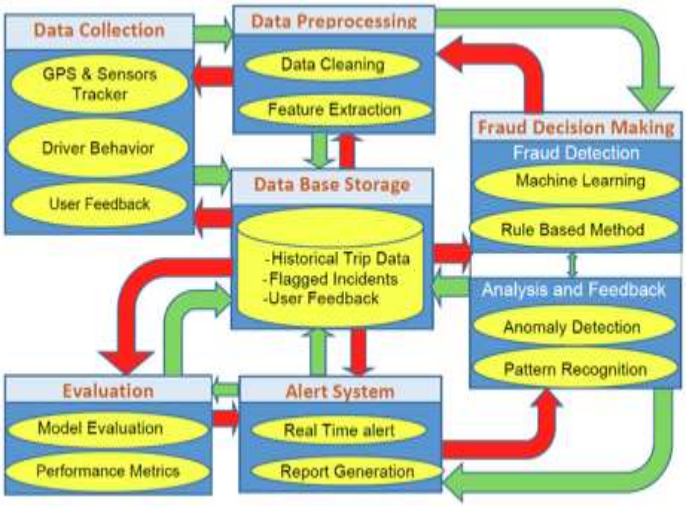

Fig.2. Architecture of a fraud taxi driver detection system

However, it is a non-trivial task to detect taxi fraud from GPS traces. There are inherent complexities involved in identifying fraudulent activities by taxi drivers. The Fig.3 highlights taxi driving fraud using GPS traces between a source (S) and a destination (E). Deviations from normal routes, such as the black trace, show excessive detours, while the red trace represents abnormally long distances. Suspicious behaviours like local detours (e.g., the blue trace) may remain undetected due to similar driving distances.

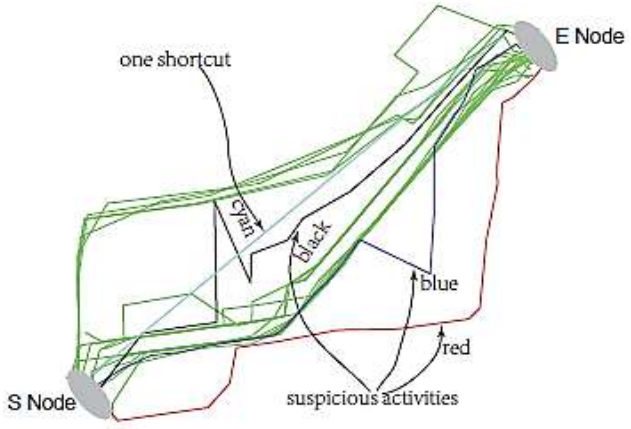

Fig.3. Taxi driving fraud through GPS traces between a source (S) and destination (E) [8]

The cyan trace illustrates a shortcut that might be misclassified as fraud by traditional anomaly detection methods. These patterns emphasize the need for more nuanced fraud detection techniques that account for spatial context and route diversity.

This review paper provides a comprehensive overview of existing approaches to identifying fraudulent taxi drivers, focusing on key phases such as pre-processing trajectory data, grouping traffic patterns, determining optimal paths, and measuring trajectory similarity. Techniques like separating vehicle data by ID, sorting locations chronologically, and map matching are explored as essential pre-processing steps. The role of machine learning is further emphasized, particularly in trajectory similarity analysis, where algorithms compare actual and optimal routes to detect deviations. Advanced techniques that include incorporation of spatial, temporal, and cost data into the predictive models towards giving more accuracy in distinguishing between fraudulent routes and the genuine ones.

This paper synthesizes the insights of previous studies through discussing strengths and limitations of the existing methods and underlines a critical role of machine learning in fraud detection systems. Spatial, temporal, and financial dimensions are identified as critical integration components in machine learning models in terms of improving accuracy and scalability. The findings presented form a basis for creating innovative, reliable, and effective solutions to tackle fraudulent behavior exhibited in the taxi industry.

The rest of the paper is arranged in five sections. The second section gives the background on key pre-processing methods of trajectory data, methods to study trajectory similarity, and finding the shortest paths. These basic concepts will be important in understanding the techniques that later will be used in the detection of fraudulent taxi behaviors. The third section talks about the literature review in order to illustrate how such methods have been applied by researchers in their different works to address the challenges in fraud detection. The forth section presents a summary of future challenges in addressing driver fraud that require more research work to find solutions for them. The last section gives a concentrated conclusion for the paper.

2. BACKGROUND AND FOUNDATION

In this section, the basic approaches used in trajectory data analysis will be summarized together with the pre-processing techniques and methods related to trajectory similarities and shortest path. These methods turn out to be very important as a means of ensuring data quality and enabling subsequent fraud detection in an effective manner.

2.1 KEY METHODS FOR PRE-PROCESSING TRAJECTORY DATA

Preprocessing trajectory data is a basic operation in the framework of managing and analyzing GPS data, aimed at cleaning the data to an accurate, reliable, and usable standard for further analysis. Here, the most relevant challenges assumed in real application scenarios involve noise, sparsity, and misalignment, are very important in fraud detection, traffic management, and route optimization. Traditional techniques and state-of-the-art machine learning/deep learning approaches are highlighted in some key methods for pre-processing trajectory data as below.

- **Map matching** is one of the most fundamental techniques to align raw trajectory data with the underlying road network. By repairing GPS inaccuracies arising from signal interference or environmental obstacles such as tall buildings and tunnels, map matching allows trajectory points to correspond to specific road segments with exactitude, thereby making possible accurate route analysis [9].
- **Noise filtering** removes artifacts or incorrect points from a trajectory, possibly caused by an error in the GPS device, a hardware malfunction, or any other environmental interferences. The main techniques used to recognize real data points from spurious ones are low pass filtering and

noise reduction by applying clustering algorithms such as DBSCAN. These techniques basically enhance the reliability of the trajectory analysis by selecting only valid data points, excluding noise [9, 10].

- **Stay point recognition** It recognizes the points that have been stationary locations where a vehicle or individual stays for a long period of time. Such points are very critical toward understanding behavioral patterns, activity locations, and points of interest. The insights drawn are valuable in transportation studies, city planning, and fraud detection [10].
- **Trajectory compression techniques** have widely investigated, aiming at reducing the volume of trajectory data with keeping of important inherent structures and key characteristics. Compression deletes the redundant or insignificant points in the track to save storage space and alleviate processing overhead without losing much analytical quality. Douglas-Peucker and spatial clustering are two of the most popular approaches here.
- **Machine learning and deep learning** have developed trajectory preprocessing by showing robust and fully automated solutions. Machine learning models can learn the pattern of data in trajectory datasets; hence, they are able to detect and correct errors in GPS automatically. Clustering algorithms, such as K-means and DBSCAN, segregate legitimate data points from anomalies effectively, thereby enhancing noise filtering capabilities. On the other hand, deep learning techniques revolutionize trajectory compression. Autoencoders reduce high-dimensional data into compact forms while retaining the critical spatial and temporal features. These models are especially helpful when there is a need to process large-scale data in real-time applications. For instance, RNNs and their variants, like LSTMs, have been explored for modeling sequential trajectory data as it enables better handling of temporal dependencies [11-13].

Besides, hybrid approaches tend to join traditional and intelligent methods. For example, map matching might be combined with ML-based outlier detection to enhance the accuracy of trajectory alignment. Similarly, combining the techniques of clustering with those of deep learning results in a more holistic approach towards trajectory pre-processing in terms of improving both anomaly detection and noise filtering respectively [13].

Moreover, to avoid noisy data, recent efforts are being made to incorporate semi-supervised learning and data augmentation techniques to reduce this dependency and improve generalization.

2.2 TRAJECTORY ANALYSIS: SIMILARITY MEASURES, SHORTEST PATHS, AND MACHINE/DEEP LEARNING

Trajectory similarity measures are quintessential tools for analyzing movement patterns and detecting fraudulent behavior within transportation systems. These measures compare trajectories to uncover deviations or irregularities that may indicate dishonest practices, such as unnecessary detours or manipulated routes. They are typically categorized into three classes: time-series, geometric, and dynamic similarity measures [14].

- **Time-series comparisons** focus on temporal patterns within trajectories. Common metrics include Euclidean Distance, Longest Common Subsequence (LCSS), Dynamic Time Warping (DTW), and Edit Distance on Real Sequences (EDR). These methods are particularly effective in fraud detection, as they analyze the timing and sequencing of a driver's route to identify inconsistencies. For example, DTW can detect whether a driver's trajectory aligns with an optimal path despite variations in speed.
- **Geometric comparisons** evaluate spatial attributes of trajectories using metrics such as Hausdorff Distance, Fréchet Distance, and Angular Trajectory Metric. These methods are especially valuable in identifying significant spatial deviations in a route. For instance, the Fréchet Distance measures how closely an actual trajectory follows an optimal path, highlighting detours or unnecessary diversions made by the driver.
- **Dynamic comparisons** assess motion characteristics such as speed and acceleration. These methods are crucial for detecting anomalies like abrupt stops or erratic driving behaviors, which may not be captured by spatial or temporal comparisons alone.

Each of these measures provides distinct insights into trajectory deviations, making them indispensable in fraud detection systems. However, their effectiveness often depends on the specific type of irregularity being analyzed. For example, geometric measures are particularly effective for identifying detours, while dynamic measures excel at detecting irregular motion patterns.

In addition to trajectory similarity measures, determining the shortest path between two points is critical for detecting fraudulent behavior. Algorithms such as Dijkstra's Algorithm, A* Search, and Floyd-Warshall Algorithm are widely employed to compute optimal routes within a network.

- **Dijkstra's Algorithm** is highly effective for identifying the most efficient route in a static network, serving as a benchmark for comparing actual driver trajectories to the ideal path.
- **A* Search** incorporates heuristics, making it particularly suitable for dynamic environments, such as real-time fraud detection systems that account for changing traffic conditions.
- **Floyd-Warshall Algorithm** calculates the shortest paths between all pairs of nodes, providing comprehensive insights into route optimization for large datasets.

These algorithms establish a baseline of "normal" routes against which actual trajectories can be compared. Deviations, such as unnecessary detours or stops, can then be flagged as potentially fraudulent. By integrating trajectory similarity measures with shortest path algorithms, fraud detection systems can effectively analyze driver behavior, uncover suspicious activities, and ensure compliance with optimal routing standards. These methods form the foundation of many modern fraud detection frameworks, enhancing accountability and trust in transportation services [15–19].

In recent developments, machine learning (ML) and deep learning (DL) methods have emerged as powerful alternatives to traditional measures in trajectory analysis. These methods leverage large-scale trajectory datasets to automatically learn patterns and identify anomalies that may indicate fraud. Examples include:

- **Clustering algorithms**, such as K-means and DBSCAN, which group similar trajectories and identify outliers that deviate significantly from the norm. This approach is particularly effective for detecting uncommon routes or suspicious driving patterns.
- **Classification models**, including decision trees and support vector machines (SVM), which are trained on labeled trajectory data to distinguish between fraudulent and legitimate behaviors.
- **Recurrent Neural Networks (RNNs)** and their variants, such as Long Short-Term Memory (LSTM) networks, which excel at modeling sequential trajectory data. These models capture temporal dependencies, enabling the prediction of expected routes and flagging deviations as potential fraud.
- **Graph Neural Networks (GNNs),** which are used to represent road networks, allowing trajectory analysis to account for spatial relationships between nodes in the network.

Machine learning and deep learning techniques complement traditional measures by offering greater adaptability and scalability, particularly in dynamic and complex scenarios where predefined similarity metrics may fall short [20]–[22].

3. LITERATURE REVIEW

One of the significant challenges facing the taxi industry is the identification of fraudulent taxi drivers. Common fraudulent activities include fare manipulation, taking unauthorized detours, and collusion with organized criminal groups. These practices result in financial losses for passengers, lack of trust in the industry, and compromise passenger well-being. Consequently, developing effective techniques for fraud detection and mitigation has become crucial. Addressing challenges within road networks, such as traffic prediction and determining the shortest path, is generally approached through two main methods: trajectory analysis and GPS stream data analysis. Below is a systematic review of recent research in the field, along with detailed comparisons.

Y. Wang et al. [23] proposed a hierarchical clustering method to analyze GPS data from taxi vehicles to identify uncommon routes and behavioral trends. This system groups similar routes to detect illogical patterns, leveraging the speed and efficiency of GPS data analysis. However, the method's accuracy depends significantly on the quality of the GPS data and struggles to detect underrepresented or subtle fraudulent behaviors in the dataset [23].

Wei et al. [24] introduced the MSD-K-means algorithm, which efficiently identifies both global and local anomalies in high-dimensional datasets. By incorporating multi-scale distance strategies, the algorithm enhances outlier detection and supports both categorical and numerical data. While the method demonstrates precision and scalability, it is highly sensitive to hyper parameters and assumptions of normal data distribution. Further optimization is needed to address issues such as closely located outliers. The study evaluated the algorithm's performance using precision, recall, and F1-score metrics, highlighting its potential for high-dimensional anomaly detection [24].

Habeeb et al. [25] conducted a comprehensive review of real-time data processing methods for anomaly detection in streaming data. The study evaluated the strengths and weaknesses of existing machine learning and deep learning systems for detecting anomalies across diverse data streams. A key advantage of these methods lies in their ability to process data in real time, but challenges such as resource-intensive operations and reliance on high-quality training data remain, requiring further research and optimization.

Yuan et al. [26] proposed a density-clustering approach for identifying anomalies in taxi trajectories. This method effectively classifies paths and detects outliers as irregular routes, enabling real-time anomaly detection in large datasets. While the system excels at identifying suspicious behavior, its accuracy depends heavily on the quality of input data and the performance of clustering algorithms, necessitating improvements for broader applicability.

Qian et al. [27] combined hierarchical clustering and decision tree algorithms to detect fraudulent taxi activities. This hybrid system is particularly effective in identifying fraud patterns such as overcharging, fare manipulation, and driver collusion, demonstrating its utility for addressing a variety of fraudulent behaviors.

Zhang et al. [28] developed a multi-step approach for anomaly detection in mobile network data, involving data preparation, feature extraction, feature selection, and classification. This method exhibits scalability and real-time anomaly detection capabilities, making it suitable for analyzing large trajectory datasets. However, its reliance on accurate data and significant resource requirements remains limitations.

Zhang et al. [29] also employed density-based clustering and feature selection methods to detect fraudulent taxi behaviors. By integrating GPS and time data, their system delivers accurate and efficient fraud detection. Nevertheless, its performance is contingent on precise and up-to-date data, highlighting the importance of continuous updates and validation.

Belhadi et al. [30] introduced a two-phase anomaly detection model that targets both individual and group outliers in ride-hailing services. Utilizing GPU acceleration, the system achieved a 341-fold speedup while maintaining high-quality anomaly detection results, demonstrating significant improvements in computational efficiency.

Karim et al. [31] explored a semantic meta-model for trajectory analysis, emphasizing the narrative aspect of fraudulent behaviors. The proposed software architecture combines graph-oriented NoSQL databases with reactive processing frameworks, enhancing fraud detection capabilities in socio-technical systems.

Kong et al. [32] presented a spatial-temporal-cost approach for fraud detection, incorporating variables such as GPS location, time, and fare data. This method demonstrated high accuracy and efficiency; however, it underscored the necessity of precise, real-time data to minimize false positives and missed detections.

Ali et al. [33] proposed a jam-distance graph-based method for optimizing route planning and shortest path determination in road networks. Leveraging Dijkstra's algorithm, the system incorporates real-time traffic conditions to improve routing accuracy. However, its effectiveness depends on the availability of accurate and up-to-date traffic data, as inaccuracies could lead to inefficient routing.

Al-Sudani et al. [34] addressed taxi fraud detection using the DBSCAN algorithm, which integrates trajectory analysis and jam-distance graphs. When evaluated with real-time GPS data, the system achieved high precision and recall scores while significantly reducing response times, demonstrating its potential to enhance urban transportation integrity.

The literature shows that recent studies have focused on detecting fraud using techniques like trajectory analysis, GPS data clustering, and anomaly detection algorithms. Methods such as hierarchical clustering for analyzing routes, density-based clustering, and hybrid machine learning models have been explored, each with its own strengths and challenges. Advanced approaches, like GPU-accelerated models and semantic meta-models, have improved efficiency and scalability but still depend on high-quality data. The review also highlights the critical role of accurate, real-time data in ensuring effective fraud detection and provides a comparative overview of the different methods and their performance metrics.

The Table.1 provides a summary of the reviewed studies, detailing their algorithms, datasets, and evaluation metrics. This summary offers a clear comparison of the strengths and limitations of each approach.

Table. 1. A Summary of Previous studies

Ref. No.	Algorithm Name	Data set	Evaluation Method
[23]	hierarchical clustering	GPS data from taxi vehicles	Precision, recall, and F1-score.
[24]	MSD-kmeans	High-dimensional datasets	Precision, recall, and F1-score
[25]	Review of methodologies	Large-scale real-time data	Comparison of existing techniques
[26]	Density Clustering	Taxi trajectories	Precision, recall, and F1-score
[27]	Hierarchical clustering and decision tree	Taxi transaction data	Not specified
[28]	Multi-step process	Extensive trajectory data	Real-world data evaluation
[29]	Density-based clustering	Taxi driver data	Evaluation trials with actual taxi data
[30]	Two-phase anomaly detection	Trajectory databases	Comparative examination
[31]	Trajectories meta-model	Socio-technical systems	Not specified
[32]	Spatial-temporal-cost	Taxi data in IoV environment	Evaluation using actual taxi driving data
[33]	Jam-distance graph	Road network	Efficiency evaluation
[34]	DBSCAN	Real-time taxi GPS data	F1-score metrics, recall, and precision

3.1 CHALLENGES AND FUTURE PROSPECTS

While considerable advancements have been made in trajectory analysis and fraud detection, critical challenges remain unresolved. Current methodologies often struggle with scalability when handling large, high-dimensional datasets generated by modern GPS systems. The reliance on static or semi-dynamic models limits the adaptability of fraud detection systems to real-time and context-aware scenarios. Moreover, trajectory similarity measures and anomaly detection models frequently face challenges in balancing computational efficiency with accuracy, especially in urban environments with complex road networks.

Future directions should focus on developing context-sensitive models that integrate multi-source data, such as weather conditions, traffic patterns, and passenger feedback, to enhance fraud detection precision. The adoption of federated learning frameworks could mitigate privacy concerns by enabling distributed model training without sharing raw data. Additionally, advancements in graph- based neural networks (GNNs) offer promising directions for representing and analyzing road networks in trajectory analysis.

Another avenue for research lies in unsupervised and semi-supervised learning techniques that reduce dependence on labeled datasets, enabling fraud detection in regions with limited data resources.

Furthermore, exploring hybrid approaches that combine deep learning with traditional optimization algorithms could address computational efficiency concerns while improving detection robustness. By leveraging these emerging technologies, future systems could achieve higher levels of accuracy, adaptability, and ethical compliance.

Another challenge is the high computational demands of deep learning and other advanced technologies. To address this, researchers have explored ways to simplify models and improve efficiency using techniques like pruning, quantization, and knowledge transfer.

Future efforts should focus also on testing fraud detection systems in real-time, measuring how quickly they respond in live environments. Using technologies like streaming data processing and edge computing can reduce delays and help detect fraud more quickly, ensuring prompt action in real-world situations.

The privacy protection is one more challenge, especially in the context of passenger tracking. This can be accomplished through strategies like data anonymization and encryption. Moreover, following regulations like the General Data Protection Regulation (GDPR) is key to safeguarding user data and supporting broader adoption. For energy challenges, studies should propose energy-efficient strategies, such as adaptive GPS sampling. This approach aims to reduce battery usage without

compromising accuracy. Future work can further optimize mobile performance through power-aware algorithms.

Current solutions often lack adaptability to the characteristics of developing regions. Future research should consider local transportation behaviors, road conditions, and informal routing patterns.

4. CONCLUSION

This review paper systematically analyzed methodologies and strategies for detecting fraudulent taxi drivers, emphasizing the critical role of trajectory analysis in addressing this issue. The study highlighted various approaches, including trajectory similarity measures, clustering techniques, and machine learning-based frameworks, demonstrating their effectiveness in identifying fraudulent behaviors. Integrating spatial, temporal, and contextual data was identified as a key factor in enhancing detection accuracy and system scalability. By synthesizing the strengths and limitations of current methods, this review provides a comprehensive foundation for researchers and practitioners. While challenges remain, such as handling complex datasets and ensuring ethical compliance, emerging technologies hold great promise for addressing these gaps. As fraud detection systems continue to evolve, leveraging advanced tools like graph-based neural networks, machine/deep learning, and hybrid models will pave the way for more robust and scalable solutions safeguarding the integrity of urban transportation systems.

REFERENCES

[1] J. Bao, P. Liu and S.V. Ukkusuri, "A Spatiotemporal Deep Learning Approach for Citywide Short-Term Crash Risk Prediction with Multi-Source Data", *Accident Analysis and Prevention*, Vol. 122, pp. 239-254, 2019.

[2] H.H. Ali, J.R. Naif and W.R. Humood, "A New Smart Home Intruder Detection System based on Deep Learning", *Al-Mustansiriyah Journal of Science*, Vol. 34, No. 2, pp. 60-69, 2023.

[3] J. Yao, Y. Ni, J. Zhao, H. Niu, S. Liu, Y. Zheng and J. Wang, "Data based Violated Behavior Analysis of Taxi Driver in Metropolis in China", *Computers, Materials and Continua*, Vol. 60, No. 3, pp. 1109-1122, 2019.

[4] S.N. AlSaad and N.M. Hussien, "Landmark based Shortest Path Detection in Alarm System", *Al-Mustansiriyah Journal of Science*, Vol. 29, No. 2, pp. 135-140, 2018.

[5] D. Zhang, N. Li, Z.H. Zhou, C. Chen, L. Sun and S. Li, "iBAT: Detecting Anomalous Taxi Trajectories from GPS Traces", *Proceedings of International Conference on Ubiquitous Computing*, pp. 99-108, 2011.

[6] C. Lou, Y. Wang, J. Li, Y. Qian and X. Li, "Graph Neural Network for Fraud Detection Via Context Encoding and Adaptive Aggregation", *Expert Systems with Applications*, Vol. 261, pp. 1-7, 2025.

[7] Y. Shen, L. Zhao and J. Fan, "Parallel Discovering of City Hot Spot based on Taxi Trajectories", *International Journal of Data Mining and Urban Planning*, Vol. 6, No. 3, pp. 215-230, 2012.

[8] Y. Ge, H. Xiong, C. Liu and Z.H. Zhou, "A Taxi Driving Fraud Detection System", *Proceedings of International Conference on Data Mining*, pp. 181-190, 2011.

[9] O. Kuhne, D. Edler and C. Jenal, "A Multi-Perspective View on Immersive Virtual Environments (IVEs)", *ISPRS International Journal of Geo-Information*, Vol. 10, No. 8, pp. 1-6, 2021.

[10] A. Korez, N. Barışçı, A. Çetin and U. Ergün, "Weighted Ensemble Object Detection with Optimized Coefficients for Remote Sensing Images", *ISPRS International Journal of Geo-Information*, Vol. 9, No. 6, pp. 1-9, 2020.

[11] J.Q. James, "Travel Mode Identification with GPS Trajectories using Wavelet Transform and Deep Learning", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 22, No. 2, pp. 1093-1103, 2020.

[12] L. Chen, M. Lv, G. Han and X. Li, "Trajectory Simplification: An Overview and Future Trends", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 21, No. 7, pp. 2921-2937, 2020.

[13] X. Zhu, X. Wang and X. Tang, "A Novel Deep Learning Approach for Trajectory Outlier Detection", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 30, No. 7, pp. 1351-1364, 2018.

[14] Y. Tao, A. Both, R.I. Silveira, K. Buchin, S. Sijben and R.S. Purves, "A Comparative Analysis of Trajectory Similarity Measures", *GIScience and Remote Sensing*, Vol. 58, No. 5, pp. 643-669, 2021.

[15] R.S.D. Sousa, A. Boukerche and A.A. Loureiro, "Vehicle Trajectory Similarity: Models, Methods and Applications", *ACM Computing Surveys*, Vol. 53, No. 5, pp. 1-32, 2020.

[16] J. Bian, D. Tian, Y. Tang and D. Tao, "Trajectory Data Classification: A Review", *ACM Transactions on Intelligent Systems and Technology*, Vol. 10, No. 4, pp. 1-34, 2019.

[17] M.W. Berry, A. Mohamed and B.W. Yap, "*Supervised and Unsupervised Learning for Data Science*", Springer, 2019.

[18] N.M. Kriege, F.D. Johansson and C. Morris, "A Survey on Graph Kernels", *Applied Network Science*, Vol. 5, No. 1, pp. 1-42, 2020.

[19] X. Liu, Z. Wang, N. Wang, X. Li, B. Zhang and J. Qiao, "An Adaptive Sharing Framework for Efficient Multi-Source Shortest Path Computation", *Proceedings of International Conference on Web Information Systems and Applications*, pp. 635-646, 2021.

[20] L. Chen, J. Liu and W. Zhou, "A Survey on Trajectory-Prediction Methods for Autonomous Driving", *IEEE Transactions on Intelligent Vehicles*, Vol. 9, No. 2, pp. 234-245, 2022.

[21] X. Li, K. Zhao, G. Cong, C.S. Jenson and W. Wei, "Deep Representation Learning for Trajectory Similarity Computation", *Proceedings of International Conference on Data Engineering*, pp. 617-628, 2020.

[22] C. Wang, J. Huang, Y. Wang, Z. Lin, X. Jin and J. Xing, "A Deep Spatiotemporal Trajectory Representation Learning Framework for Clustering", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 23, No. 10, pp. 7687-7700, 2024.

[23] Y. Wang, K. Qin, Y. Chen and P. Zhao, "Detecting Anomalous Trajectories and Behavior Patterns using Hierarchical Clustering from Taxi GPS Data", *ISPRS*

International Journal of Geo-Information, Vol. 7, No. 1, pp. 1-7, 2018.

[24] Y. Wei, J. Jang-Jaccard, F. Sabrina and T. McIntosh, "MSD-kmeans: A Novel Algorithm for Efficient Detection of Global and Local Outliers", *Proceedings of International Conference on Machine Learning*, pp. 1-6, 2019.

[25] R.A.A. Habeeb, F. Nasaruddin, A. Gani, I.A.T. Hashem, E. Ahmed and M. Imran, "Real-Time Big Data Processing for Anomaly Detection: A Survey", *International Journal of Information Management*, Vol. 45, pp. 289-307, 2019.

[26] H.U. Yuan, L.I. Hui and C.H.E.N. Mei, "Taxi Abnormal Trajectory Detection based on Density Clustering", *Computer and Modernization*, Vol. 5, No. 6, pp. 1-5, 2019.

[27] S. Qian, B. Cheng, J. Cao, G. Xue, Y. Zhu and J. Yu, "Detecting Taxi Trajectory Anomaly based on Spatio-Temporal Relations", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 23, No. 7, pp. 6883-6894, 2021.

[28] H. Zhang, Y. Luo, Q. Yu, L. Sun, X. Li and Z. Sun, "A Framework of Abnormal Behavior Detection and Classification based on Big Trajectory Data for Mobile Networks", *Security and Communication Networks*, Vol. 2020, No. 4, pp. 1-15, 2020.

[29] Z. Zhang, M. Li, F. He and Y. Wang, "Clustering Approach for Trajectory Anomaly Detection", *Proceedings of International Conference on Transportation Professionals*, pp. 113-124, 2020.

[30] A. Belhadi, Y. Djenouri, G. Srivastava, D. Djenouri, A. Cano and J.C.W. Lin, "A Two-Phase Anomaly Detection Model for Secure Intelligent Transportation Ride-Hailing Trajectories", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 22, No. 7, pp. 4496-4506, 2020.

[31] L. Karim and A. Boulmakoul, "Trajectory-based Modeling for Fraud Detection and Analytics: Foundation and Design", *Proceedings of International Conference on Computer Systems and Applications*, pp. 1-7, 2021.

[32] X. Kong, B. Zhu, G. Shen, T.C. Workneh, Z. Ji, Y. Chen and Z. Liu, "Spatial-Temporal-Cost Combination based Taxi Driving Fraud Detection for Collaborative Internet of Vehicles", *IEEE Transactions on Industrial Informatics*, Vol. 18, No. 5, pp. 3426-3436, 2021.

[33] S.F. Ali, M.R. Abdulrazzaq and M.T. Gaata, "Finding Shortest Path in Road Networks based on Jam-Distance Graph and Dijkstra's Algorithm", *Proceedings of International Conference on Next Generation of Internet of Things*, pp. 469-480, 2022.

[34] Z.S. Al-Sudani and M. Riyadh, "Fraudulent Taxi Driver Detection in Baghdad City based on DBSCAN and A* Algorithm", *Proceedings of International Conference on Communication and Information Technology*, pp. 165-170, 2023.