Reyes, G., Lanzarini, L., Estrebou, C., Bariviera, A., Maquilón, V. (2023). Evaluation of a Grid for the Identification of Traffic Congestion Patterns. In: Valencia-García, R., Bucaram-Leverone, M., Del Cioppo-Morstadt, J., Vera-Lucio, N., Centanaro-Quiroz, P.H. (eds) Technologies and Innovation. CITI 2023. Communications in Computer and Information Science, vol 1873. Springer, Cham. doi:10.1007/978-3-031-45682-4_20

Evaluation of a Grid for the Identification of Traffic Congestion Patterns

Gary Reyes[1], Laura Lanzarini[2], César Estrebou[2], Aurelio F. Bariviera[3], and Victor Maquilón[1]

[1] Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil, Cdla. Universitaria Salvador Allende, Guayaquil 090514, Ecuador
[2] Universidad Nacional de La Plata, Facultad de Informática, Instituto de Investigación en Informática LIDI (Centro CICPBA), 1900, La Plata, Buenos Aires, Argentina
[3] Department of Business, Universitat Rovira i Virgili, Reus, Spain

Abstract. Today, urban growth, increased vehicular traffic and congestion have become a key challenge in cities. As a consequence, negative effects on mobility are generated, such as longer travel times, increased environmental pollution, stress for drivers, and difficulties in urban traffic planning and management. Understanding and analyzing congestion patterns is essential to effectively address this problem and develop more efficient traffic management strategies. Some research has proposed various solutions to address vehicular congestion, such as the use of algorithms for traffic data analysis, the implementation of intelligent traffic management systems, and the optimization of road infrastructure. The proposed methodology uses dynamic clustering techniques and the analysis of historical information to analyze vehicular congestion patterns, implementing the DyClee algorithm adapted to cells. The obtained results on the city of San Francisco are satisfactory, allowing the identification of clusters with certain patterns that allow identifying areas and times of higher congestion, revealing the temporal variability and highlighting the importance of considering the dynamics of vehicular flow in traffic management.

Keywords: Dynamic clustering · Data stream · Congestion · Grid

1 Introduction

Efficient traffic management is a constant challenge in urban areas, where vehicle flow can be intense and congested. Accurate understanding and analysis of vehicular flow is essential to improve transportation planning and road infrastructure decision making. In this context, this paper presents a methodology for analyzing vehicular flow based on the analysis of vehicle trajectory data.

The study of vehicular flow has been approached from different approaches, and one of the key aspects lies in the adequate representation of the trajectory data. In this sense, the proposed methodology is based on the division of a geographical area of interest into uniform cells, where each cell contains detailed information on the speed and number of vehicles. This representation simplifies the analysis and visualization of the data, allowing a more accurate understanding of the vehicular flow in the study area.

A clustering algorithm is used to perform the traffic flow analysis. It is based on the methodology proposed by Reyes et al. [17]. This algorithm focuses on efficient and accurate clustering of cells containing vehicle trajectories. Using a distance metric that considers the minimum difference in speed and number of vehicles between the groups and cells being analyzed, the algorithm optimally groups the cells. This allows the identification of patterns and similarities in the vehicular flow of each group, facilitating the detection of areas with potential congestion problems.

The evaluation of vehicular flow congestion is performed using a dataset of historical trajectories. The characteristics of the cells belonging to the microclusters resulting from the clustering are analyzed and compared with reference values obtained from adjacent cells. Through a logical evaluation, it is determined whether the cells meet the established congestion conditions. This evaluation provides a clear view of the areas with traffic flow problems, providing relevant information for traffic management decisions.

This article is organized as follows: Sect. 2 analyzes some related works that were identified within the literature and present various solutions to the problem, Sect. 3 describes the evaluation of the proposed method, Sect. 4 discusses the obtained results, and Sect. 5 presents the conclusions and future lines of work.

2 Related Work

The use of clustering techniques has been extensively explored in the context of data flow analysis [7, 20]. In the field of trajectory analysis, a wide variety of research has contributed to the development of clustering techniques applied to different domains [14]. These works have addressed the need to adapt conventional algorithms to handle trajectories.

Clustering methods have been used to identify groups with similar characteristics and analyze their collective behavior [5], this database has allowed the identification of patterns and structures in the vehicle trajectory data, which in turn has facilitated the detection of areas with similar characteristics to traffic congestion.

Within the literature, research can be found that has addressed trajectory segmentation as a stage prior to the clustering process. For example, Lee et al. [10] have proposed an approach that uses segmentation to improve the performance of clustering algorithms and obtain higher quality clusters. Similarly, Mao et al. [13] & Yuan et al. [22] also presented methods that incorporate trajectory segmentation prior to clustering, which has proven beneficial for trajectory analysis.

Some approaches have used traditional clustering algorithms, such as the k-means algorithm [7] and the DBSCAN algorithm [3], adapted specifically for path analysis.

These methods consider similarity metrics specially designed for trajectories, allowing them to efficiently cluster and classify according to desired characteristics and appropriate situations [1]. Researchers have proposed several similarity measures. For example, the Tra-DBScan algorithm [11] adapted the well-known DBScan algorithm [3] and introduced the Hausdorff distance as a similarity measure to divide trajectories into sections. Reyes et al. [18] proposed a pivot-guided similarity metric that uses angular information for GPS vehicle trajectory segmentation. In addition, Yu et al. [21] developed a new clustering algorithm that calculates the similarity between two trajectories using multiple data features.

Other approaches have explored the use of dynamic clustering algorithms, which can handle the continuous flow of trajectory data and adapt to changes in traffic behavior over time. These dynamic approaches allow early detection of changes in vehicular flow and identification of areas with emerging congestion [12].

In addition to traditional path clustering techniques, approaches based on complex network analysis and graph theory have also been explored to understand vehicular flow dynamics.

In the related literature, different approaches have been proposed for real-time vehicular flow analysis. Some researchers have used machine learning [8] and data mining techniques, such as neural networks and regression models, to identify hidden patterns and anomalies in traffic flow [19]; clustering and classification techniques [2, 23] have also been applied to analyze real-time traffic data and detect congestion on urban roads. These approaches rely on the collection of real-time data, such as traffic signals and GPS data, to train models that can make accurate forecasts. Other researchers have used time series models to predict vehicular flow and take proactive measures in traffic management [24]. These works demonstrate the importance of combining clustering techniques with other analysis methodologies to obtain a complete picture of vehicular flow under different scenarios.

These approaches based on complex networks and data mining offer alternative perspectives to understand vehicular flow from a systemic and predictive point of view [9]. While approaches based on the integration of multiple data sources offer opportunities to improve the understanding of vehicular flow and optimize real-time traffic management [12].

The use of clustering techniques in vehicular flow analysis is an active field of research, with continuous advances in improving the accuracy and efficiency of clustering algorithms, which in turn contributes to better traffic management and informed decision making in real time. The paper proposed by Reyes et al. [16] presents a methodology capable of analyzing vehicular flow in a given area, identifying speed ranges and maintaining an updated interactive map that facilitates the identification of areas with potential traffic jams.

However, a key issue that has been identified in these studies is the need for complementary datasets to enrich the analysis of vehicular flow [4]. In this sense, the methodology presented in this article addresses this need by incorporating an additional dataset of historical trajectories, allowing us to obtain a more complete and accurate perspective of vehicle behavior in the area of interest.

The methodology used makes use of a grid for the representation of vehicular flow, a dynamic grouping of cells and a specific analysis of each cell to determine if it meets certain congestion conditions based on reference values of speed and number of vehicles. This approach provides an accurate comparative basis and allows an assessment of congestion conditions in the study area. By combining the existing approaches with an additional set of historical trajectory data and grid representation, a more complete and accurate analysis of vehicular flow over the clustering results in the area of interest, thus allowing for early and accurate detection of areas with potential traffic flow problems.

3 Methodology

This paper presents a cell evaluation methodology for the identification of congestion patterns, the methodology proposed in this paper is based on the methodology presented by Reyes et al. [16] using a dynamic clustering algorithm, data representation by cells and visualization of clustering results. However, additional elements are incorporated, such as an additional dimension for clustering, the use of a complementary dataset and the evaluation of congestion features in the clusters.

First, the definition of the cells to be analyzed is carried out. The second step involves the use of a proposed dynamic clustering algorithm in order to identify patterns of areas that could present congestion. In the third step, the identified patterns are analyzed and evaluated with a congestion feature evaluation method. Then, in the fourth step, we proceed to generate an appropriate visualization for the obtained results. Each of these steps is described in detail below.

3.1 Definition of Cells

Within the structure of the records to be processed are the most important attributes such as longitude, latitude, time, speed and the trajectory identifier.

The first crucial step in the methodology is to ensure an adequate representation of the data that make up the trajectories. To achieve this, we start by defining the area of interest, thus delimiting the geographic region to which the trajectories to be analyzed belong. Once the area has been established, we proceed to subdivide it into uniform cells or smaller zones. The size of each cell is determined according to the required precision for the analysis of the vehicular flow, allowing us to capture the most relevant details of the movement patterns.

Several experiments have been conducted with the objective of determining the appropriate cell size for vehicular flow analysis. These experiments prior to the ideal parameterization have been fundamental to select the optimal configuration that guarantees an adequate representation of the movement patterns in the area of interest.

In this work, cells of 30×30 m were used and the essential characteristics of these cells by which they will be clustered will be the speed of the cells with ranges between 20 km/h and the number of vehicles with ranges between 5 vehicles.

This aspect acquires a fundamental relevance, since it implies the understanding of the information corresponding to each cell during a specific time interval. The proposed methodology in this study consists of integrally analyzing the events in each cell instead of addressing the trajectories of individual vehicles. This approach simplifies both the analysis and the visualization of the data.

Specifically, in this paper, the related data to vehicular flow, represented in each cell, were subjected to analysis using a microbatch approach in uniform time periods. The duration of each period is an essential parameter of the algorithm, which must be set in a predefined manner.

Each period is referred to as a cycle and represents an evolution of the formed clusters, as the results of the grouping are updated by incorporating each block of data. This methodology allows capturing the dynamics progressively, as new data is integrated, generating a more accurate and updated understanding of the vehicular flow.

In each evolution cycle, a continuous flow of data is stored in a temporary buffer and the relevant calculations are performed for each cell, allowing to obtain characteristic and relevant information.

3.2 Dynamic Cell Clustering

In the second step of the methodology, an algorithm for processing cells containing trajectories is used. This algorithm has been selected for its ability to efficiently and accurately analyze the specific characteristics of the trajectories in each cell.

As mentioned above, the first stage of the DyClee algorithm is used to construct microclusters. However, in this paper, an adaptation of the original proposed methodology has been made, where instead of using directly the GPS locations and their density, the speeds of the trajectory sections and the number of vehicles present in each cell are considered for each cycle.

At the stage of grouping the cells in each cycle, a distance metric is used that considers the minimum difference in speed and number of vehicles between the microclusters and the analyzed cell. If the difference in speed and vehicles is sufficiently small, the cell is incorporated into the existing microcluster and the corresponding information is updated. If there is no nearby cluster, a new microcluster is created with the analyzed cell. In this way, an appropriate grouping of cells based on similarities in speed and number of vehicles is achieved.

In the second stage of the algorithm, an analysis of the densities of the microclusters is carried out, classifying them into two categories: dense and not very dense. The number of formed microclusters is related to the value of the "relative size" parameter, which directly modifies the size of the "hyperboxes". This is reflected in a greater or lesser amplitude of the range of cell assignment to the microclusters, according to their speed and number of vehicles.

During these processes, information regarding the performance of the stages is recorded separately, allowing a detailed analysis of their operation.

3.3 Congestion Evaluation

In this study, it is important to have a dataset of historical trajectories that differs in date but matches the time and location of the one used in the clustering process. This historical data set corresponds to a real data flow, provides complementary information and enriches the analysis of vehicle flow in transit. This ensures a more complete and accurate perspective of the behavior of vehicles in the cells of the area of interest.

For the representation of the vehicular flow, a projection of a grid similar to the one used in Sect. 3.1 of the article is used, this grid is applied to the historical data set. Then, a specific analysis is performed to determine a reference value for speed and number of vehicles. This reference value is obtained by considering adjacent cells and using up to five levels of distance, these historical reference values are calculated proportionally where each level contributes a specific proportion.

As shown in Fig. 1, the cells on the periphery will be determined with lower proportion values. Being level 1, the cell being analyzed, represents 30% of the total historical value, decreasing in percentage according to the level, up to level 5, which contributes 10% of the total. In this way, information from nearby cells is used to establish a more accurate comparative basis.

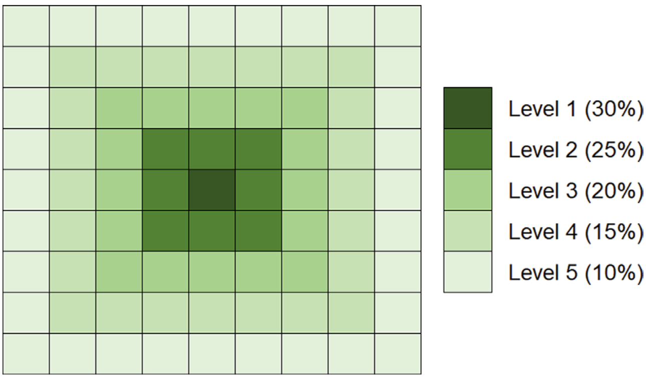

Fig. 1. Proportions according to cell level

Subsequently, the characteristics of the cells in the groups resulting from the grouping are compared with the reference values obtained previously. This comparison makes it possible to evaluate whether the analyzed cells comply with the established congestion conditions.

The conditions for congestion detection are based on the proposed evaluation by He et al. [6]. A logical evaluation is applied to determine whether the analyzed cells meet the following congestion conditions:

- The number of path points in the cell is greater than or equal to the number of path points set as a reference.
- The average speed of the vehicles inside the cell is lower than the established reference speed.

This logical evaluation makes it possible to identify the cells that show signs of congestion and contributes to the early and accurate detection of areas with potential traffic flow problems.

To provide an overview of the performed evaluations, the congestion index is calculated using the formula proposed by He et al. [6]. This defines the index as the total number of cycles for which each cell has complied with the congestion conditions set forth indicating that they have been identified as areas of congestion.

3.4 Clustering Visualization

In this study, the trajectory information is analyzed in 3-minute intervals, this time allows identifying changes in the vehicular flow in a more accurate and detailed way.

In order to provide an interactive visual representation of the results of each cluster, an interactive map has been developed. This map allows the relevant information for each cluster to be analyzed graphically and dynamically. Each cluster is represented on the map with a different color for each pattern based on the number of vehicles and speeds that the clustering algorithm has identified, as shown in Fig. 2.

This study records the result of the grouping process performed, which allows reconstructing all the maps from the beginning of the vehicular flow analysis. This provides a quick overview of the traffic status.

4 Results and Discussion

4.1 Used Data

To test the performance of the proposed methodology in this article, a dataset collected in the city of San Francisco was used. A brief description of the data set is provided below:

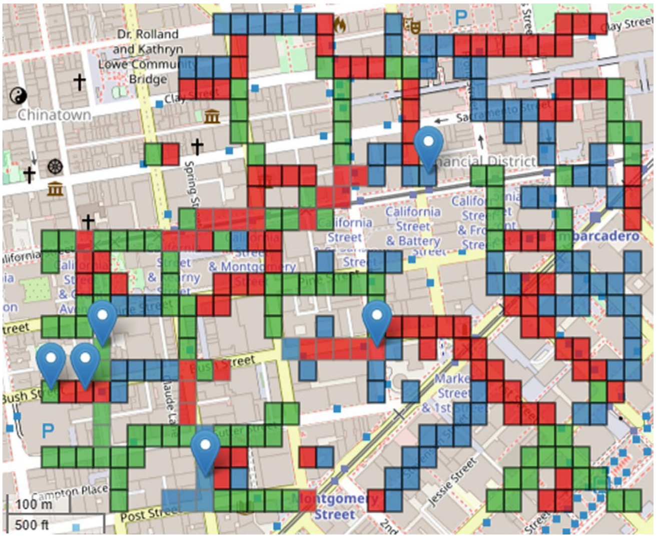

Fig. 2. Clusters projected on the map

San Francisco Dataset. A portion of the san francisco dataset [15] was extracted for clustering, the data corresponds to June 2, 2008; it contains 290 trajectories recorded by cabs using GPS devices. Each record contains the following data: trajectory id, latitude, longitude, time, speed and direction. For this set of trajectories, the analysis included all recorded trajectories between 12:30 pm and 13:30 pm. As a result of this filtering process, 191,315 records were obtained, representing 290 trajectories from the entire dataset.

4.2 Obtained Results

To perform the analysis, the dataset was divided into cycles covering 3 min of data, which generated a total of 20 cycles for the city of San Francisco. These blocks were analyzed consecutively, considering that the 3-minute period is adequate in relation to the volume of data available.

Next, we proceeded to define the values of the “relative size” and “hyperbox” parameters to use the DyClee algorithm. A “relative size” value of 0.2 was set, with speed limits between 0 and 120 km/h, and vehicle quantity limits between 1 and 25 for each cell.

Figure 3 shows the projection of the cells on a plane equivalent to the map of the processing area analyzed as a result of the evaluations according to the referential values of each location performed on the 20 cycles of the data set. Each congested cell is shown in red color indicating that these cells have met the two conditions stated in Sect. 3.3 which indicate that the cells have been evaluated as congested.

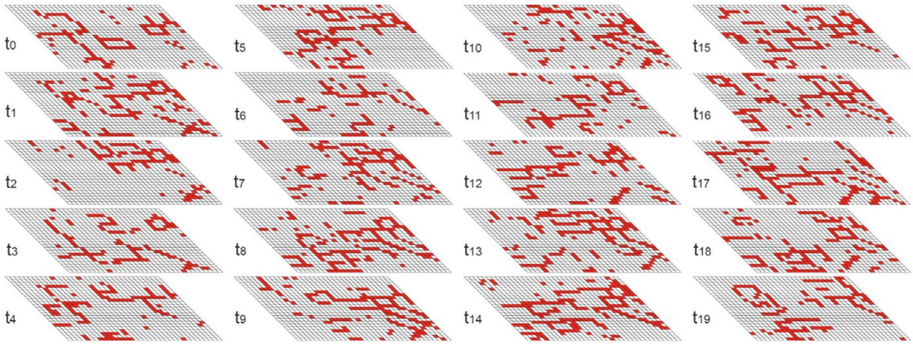

Fig. 3. Evaluation of congested cells using historical averages

Each map shows the locations where the congestion conditions are met, it can be seen that these vary over time and are different depending on the location. These variations can range from concentrations in specific zones (such as Fig. 3 - $\boldsymbol{t}_2$), zones with dispersed congestion (such as Fig. 3 - $\boldsymbol{t}_{10}$), zones with congestion in peripheral areas (such as Fig. 3 - $\boldsymbol{t}_{18}$).

These patterns of variation in congestion conditions can be attributed to various factors, such as time of day, special events or even weather conditions.

It can also be observed that at certain times, clusters of congested cells are identified that form specific spatial patterns. For example, in some cycles, congestion "spots" are observed to form in specific areas such as interconnected roads that are very busy, while in other cycles, more diffuse and dispersed congestion can be identified indicating that congestion may be occurring due to particular situations. These spatial patterns may be related to the road infrastructure, the distribution of points of interest or even the population density in different areas.

It is important to note that not all evaluated cells show congestion in all cycles. Some cells may alternate between congested and non-congested conditions over time, indicating the dynamicity and variability related to vehicular flow.

The experiments and analysis of the proposed methodology were carried out using a computer with the following technical specifications: AMD A9-9425 3.1 GHz processor; 8 GB DDR4 RAM; and Windows 10 Home operating system.

Table 1. Execution times

Description	Clustering (Sec.)	Evaluation (Sec.)	Total (Sec.)
Total time	2458.218	13.333	2471.551
Average cycle time	122.911	0.667	123.578
Deviation	±39.410	±0.124	±39.534

The obtained results about the performance of the algorithm are shown in the Table 1, this table contains a summary of the execution times of the proposed methodology. The clustering time used 2458.218 s and the evaluation of the cells took 13.333 s, which adds up to a total time of 2471.551 s.

Analyzing the average time per cycle, it is observed that the clustering requires on average 122.911 s with a deviation of ±39.410 s, while the evaluation only needs 0.667 s with a deviation of ±0.124 s. These data indicate that the clustering process is more time demanding compared to evaluation. On average, each cycle of the data set takes approximately 123.578 s to complete. These results reflect the effectiveness of the methodology in terms of execution times.

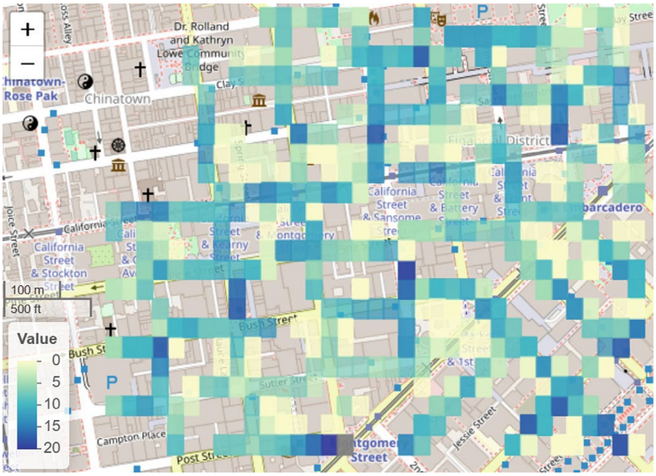

Fig. 4. Congestion distribution by historical averages

To provide an overview of the assessments made and to quantitatively represent the vehicular congestion in the analyzed processing area, the congestion index is calculated which provides a single measure for each cell and is based on the accumulation of the cells identified as congested over the 20 cycles of the data set.

The result of this calculation provides a numerical value representing the level of congestion in each cell. The higher the value of the congestion index, the higher the frequency of vehicular congestion at that specific location. This approach provides a global perspective of congestion in the processing area and facilitates the identification of critical areas requiring attention and possible solutions.

For the visualization of these indices, a congestion distribution is plotted showing a spatial representation of the areas with higher or lower frequency of congestion, which helps to understand the geographic distribution of this phenomenon. This distribution is shown in Fig. 4.

In order to understand the relationship between density and congestion classification, a contingency matrix has been used and is shown in Table 2. Analysis of this contingency matrix reveals trends in accuracy of congestion rankings relative to the density of the resulting clusters.

Table 2. Contingency matrix based on matching the results of pooling and actual flow evaluation

Categories	Congested	Non-congested
Low dense	60.50%	57.60%
Dense	22.31%	60.68%

These values represent the percentage of correctly classified cells in each combination of categories. In the category of sparse groups, resulting from clustering, 60.50% of the congested cells match cells classified as congested from the actual cell flow, and 57.60% of the uncongested cells match cells classified as uncongested from the actual flow. Likewise, in the dense group category, 22.31% of the congested cells match with cells classified as congested from the actual cell flow, and 60.68% of the uncongested cells match with cells classified as uncongested from the actual flow.

It is observed that the accuracy of the classifications varies according to the density category resulting from the clustering. This indicates that the classification result performs better in identifying congestion in clusters with low cell density. By understanding the relationship between traffic density and congestion classifications, more effective classification strategies can be developed for accurate identification of vehicular congestion in variable traffic environments.

In order to improve the congestion detection match rates using the proposed methodology, it is crucial to consider several factors. First, the accuracy of the traffic data must be ensured by obtaining high quality and up-to-date information from reliable sources. In addition, the parameters of the clustering algorithm and congestion thresholds need to be reviewed and adjusted to suit the specific characteristics of the city or study area. The inclusion of external factors, such as special events or weather changes, is also relevant to have a complete understanding of congestion and can significantly improve the accuracy of vehicle congestion detection.

The methodology developed in this study presents similarities with the one proposed by He et al. [6] for traffic congestion event detection processes. Although there are differences in the implementation details and specific objectives, both methodologies share a common conceptual basis to address the identification of vehicular congestion. This similarity allows the possibility of making partial comparisons between both methodologies and enriching the general knowledge on traffic congestion detection in dynamic environments.

It is important to recognize that the proposed methodology also faces certain limitations and potential challenges that must be taken into account when interpreting the results. Since congestion conditions can vary widely depending on time of day, special events or even weather conditions, this methodology may generate different results at different times. Careful analysis of the results in different temporal contexts is essential to understand the stability of the identified patterns.

In addition, although the methodology has proven to be effective in the city of San Francisco, results should be validated in other urban areas with different characteristics. Each city may have unique roadway infrastructures, distinct traffic patterns, and particular congestion dynamics. This may require specific adjustments and adaptations to obtain accurate and meaningful results.

Another potential limitation is the need for high quality data. The success of the methodology is highly dependent on the availability of accurate vehicular traffic data. If data are insufficient or of low quality, this could affect the accuracy and reliability of congestion assessments.

5 Conclusions

The distribution of vehicular congestion in the study area has been clearly identified and visualized, allowing a better understanding of the spatial and temporal patterns of vehicular congestion.

The effectiveness of the proposed methodology for accurately and systematically assessing and monitoring vehicular congestion has been demonstrated.

The obtained results have provided valuable information on the areas and times of greatest congestion, which can contribute to informed decision making in urban traffic planning and management.

Significant variability in congestion conditions over time has been observed, highlighting the importance of considering the seasonality and dynamics of vehicular flow in traffic management strategies.

Visualization of the results has made it possible to identify specific patterns of congestion in specific areas, which can guide actions aimed at improving road infrastructure or implementing specific traffic control measures in those locations.

The methodology has proven to be an effective tool for the identification and monitoring of congested areas, which can contribute to the optimization of resources and efficiency in urban transport management.

The obtained results support the need to implement traffic control and urban planning strategies that take into account the variability and dynamics of vehicular congestion, with the objective of improving fluidity and reducing travel times in the study area.

As lines of future work, it is proposed to improve the classification model to improve the accuracy in the classification of congested and non-congested cells. Another proposal is to consider additional variables such as vehicle flow speed, roadway capacity or type of vehicles to obtain a more complete understanding of congestion. Evaluating vehicle congestion over time to develop predictive models to identify long-term patterns and trends. In addition, validation of the new methodology in other spatial and temporal contexts is proposed to verify its applicability and generalizability. This would allow the new methodology to be used to improve traffic management in a variety of environments.

References

1. Choong, M.Y., Chin, R.K.Y., Yeo, K.B., Teo, K.T.K.: Trajectory pattern mining via clustering based on similarity function for transportation surveillance. Int. J. Simul. Syst. Sci. Technol. **17**(34), 1–19 (2016)
2. Erdelić, T., Carić, T., Erdelić, M., Tišljarić, L., Turković, A., Jelušić, N.: Estimating congestion zones and travel time indexes based on the floating car data. Comput. Environ. Urban Syst. **87**, 101604 (2021). https://doi.org/10.1016/j.compenvurbsys.2021.101604. https://www.sciencedirect.com/science/article/pii/S0198971521000119
3. Ester, M., Kriegel, H.P., Sander, J., Xu, X., et al.: A density-based algorithm for discovering clusters in large spatial databases with noise. In: Kdd, vol. 96, pp. 226–231 (1996)
4. Gao, H., et al.: A method for exploring and analyzing spatiotemporal patterns of traffic congestion in expressway networks based on origin-destination data. ISPRS Int. J. Geo-Inf. **10**(5) (2021). https://doi.org/10.3390/ijgi10050288. https://www. mdpi.com/2220-9964/10/5/288
5. Han, J., Kamber, M., Tung, A.K.: Spatial clustering methods in data mining. In: Geographic Data Mining and Knowledge Discovery, pp. 188–217 (2001)
6. He, Y., Hofer, B., Sheng, Y., Yin, Y., Lin, H.: Processes and events in the cen-ter: a taxi trajectory-based approach to detecting traffic congestion and ana-lyzing its causes. Int. J. Digit. Earth **16**(1), 509–531 (2023). https://doi.org/10.1080/17538947.2023.2182374. https://www.tandfonline.com/doi/full/10.1080/17538947.2023.2182374
7. Jain, A.: Data clustering: 50 years beyond k-means. Pattern Recogn. Lett. **39**, 651–666 (2009)
8. Kamble, S.J., Kounte, M.R.: Machine learning approach on traffic congestion mon-itoring system in internet of vehicles. Procedia Comput. Sci. **171**, 2235–2241 (2020). https://doi.org/10.1016/j.procs.2020.04.241. https://www.sciencedirect.com/science/article/pii/S1877050920312321. Third International Conference on Computing and Network Communications (CoCoNet 2019)
9. Kim, J., Mahmassani, H.S.: Spatial and temporal characterization of travel pat-terns in a traffic network using vehicle trajectories. Transp. Res. Procedia **9**, 164–184 (2015)
10. Lee, J.G., Han, J., Whang, K.Y.: Trajectory clustering: a partition-and-group framework. In: Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data - SIGMOD 2007, p. 593. ACM Press (2007). https://doi.org/10.1145/1247480.1247546. http://portal.acm.org/citation.cfm?doid=1247480. 1247546
11. Liu, L.X., Song, J.T., Guan, B., Wu, Z.X., He, K.J.: Tra-DBScan: a algorithm of clustering trajectories. In: Applied Mechanics and Materials, vol. 121, pp. 4875–4879. Trans Tech Publ (2012)
12. Lou, J., Cheng, A.: Detecting pattern changes in individual travel behavior from vehicle GPS/GNSS data. Sensors **20**(8) (2020). https://doi.org/10.3390/ s20082295. https://www.mdpi.com/1424-8220/20/8/2295
13. Mao, Y., Zhong, H., Qi, H., Ping, P., Li, X.: An adaptive trajectory clustering method based on grid and density in mobile pattern analysis. Sensors **17**(9), 2013 (2017). https://doi.org/10.3390/s17092013. http://www.mdpi.com/ 1424-8220/17/9/2013

14. Mazimpaka, J.D., Timpf, S.: Trajectory data mining: a review of methods and applications. J. Spat. Inf. Sci. **2016**(13), 61–99 (2016)
15. Piorkowski, M., Sarafijanovic-Djukic, N., Grossglauser, M.: CRAWDAD dataset EPFL/mobility (v. 2009–02-24) (2009). https://doi.org/10.15783/C7J010
16. Reyes, G., Lanzarini, L., Estrebou, C., Fernandez Bariviera, A.: Dynamic grouping of vehicle trajectories. J. Comput. Sci. Technol. **22**(2), e11 (2022). https://doi.org/10.24215/16666038.22.e11. https://journal.info.unlp.edu.ar/JCST/article/view/2059
17. Reyes, G., Lanzarini, L., Estrebou, C., Maquilón, V.: CACIC 2021 XXVII CONGRESO ARGENTINO DE CIENCIAS DE LA COMPUTACION, 2021 edn., vol. XXVII, p. 261–270. Universidad Nacional de Salta (2021)
18. Reyes-Zambrano, G., Lanzarini, L., Hasperué, W., Bariviera, A.F.: GPS trajectory clustering method for decision making on intelligent transportation systems. J. Intell. Fuzzy Syst. **38**(5), 5529–5535 (2020). https://doi.org/10.3233/JIFS-179644
19. Sun, S., Chen, J., Sun, J.: Traffic congestion prediction based on GPS trajectory data. Int. J. Distrib. Sensor Netw. **15**(5), 1550147719847440 (2019). https://doi.org/10.1177/1550147719847440. Publisher: SAGE Publications
20. Tork, H.F.: Spatio-temporal clustering methods classification. In: Doctoral Sym-posium on Informatics Engineering, vol. 1, pp. 199–209. Faculdade de Engenharia da Universidade do Porto Porto, Portugal (2012)
21. Yu, Q., Luo, Y., Chen, C., Chen, S.: Trajectory similarity clustering based on multi-feature distance measurement. Appl. Intell. **49**(6), 2315–2338 (2019)
22. Yuan, G., Sun, P., Zhao, J., Li, D., Wang, C.: A review of moving object trajectory clustering algorithms. Artif. Intell. Rev. **47**(1), 123–144 (2017). https://doi.org/10.1007/s10462-016-9477-7. http://link.springer.com/10.1007/s10462-016-9477-7
23. Zhang, Y., Ye, N., Wang, R., Malekian, R.: A method for traffic congestion cluster-ing judgment based on grey relational analysis. ISPRS Int. J. Geo-Inf. **5**(5) (2016). https://doi.org/10.3390/ijgi5050071. https://www.mdpi.com/2220-9964/5/5/71
24. Zhou, R., Chen, H., Chen, H., Liu, E., Jiang, S.: Research on traffic situation anal-ysis for urban road network through spatiotemporal data mining: a case study of xi'an, china. IEEE Access **9**, 75553–75567 (2021). https://doi.org/10.1109/ ACCESS.2021.3082188