Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2026) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

World Conference of Transport Research, Toulouse 2026 (WCTR 2026 Toulouse)

A Time-Window GNN-Based Network Partition for Identifying High- and Low-Risk Nodes in Road Networks

Simone Paradiso[a, b*], Apostolos Ziakopoulos[a], Haris Sideris[b], George Yannis[a]

[a]*National Technical University of Athens, Department of Transportation Planning and Engineering, 5 Iroon Polytechneiou Street, GR-15773, Athens, Greece*
[b]*OSeven Telematics, 27B Chaimanta Street, 15234 Chalandri, Greece*

Abstract

Road safety analysis aims to reduce road crashes and enhance transportation systems' safety and efficiency. Spatial analysis of road networks plays a crucial role in evaluating and understanding these systems and how they function across different spatial localities. To that end, Artificial Intelligence (AI) tools, such as Graph Neural Networks (GNN), provide a robust framework for analyzing data structured as networks, where nodes and edges represent spatial entities. The present study exploits telematics data from naturalistic driving, which include surrogate safety measures such as harsh braking and harsh acceleration indicators, grouped by weekly time windows and integrated into a road network enriched with geometric features. This telematics-informed network is then used as input to a GNN, which generates node embeddings, which are clustered to produce weekly node-based road network partitions, revealing spatiotemporal patterns and enabling proactive identification of unsafe areas. Compared to directly clustering the raw telematics-informed network, this GNN-based approach yields better internal validation scores, improving cluster interpretability and reliability, and greater consistency across time windows. This approach enables proactive road safety management by uncovering evolving spatial risk patterns over time, supporting more targeted and detailed interventions.

Keywords: Spatial Road Safety Analysis; Graph Neural Networks; Network Partitioning; Time-window Node Embeddings

1. Introduction

Despite some progress, road safety remains a critical global issue, as there were an estimated 1.19 million road traffic deaths in 2021, a 5% drop when compared to the 1.25 million deaths in 2010. Road crashes are still the leading cause of death of children and youth, and they typically occur during peoples' most productive years, causing huge health, social and economic harm (WHO, 2023).

Road transport involves distances, making spatial analysis a practical approach when examining crashes. Econometric models have been widely used to predict crashes and their severity, and with the advent of big data, the use of Machine Learning (ML) and Artificial Intelligence (AI) has been rapidly increased in spatial road safety analysis (Ziakopoulos & Yannis, 2020). Along with these, Deep Learning (DL), based on Artificial Neural Networks (ANNs), has also gained popularity; these models use layers of computational units, called neurons, that learn non-linear input-output mappings allowing the model capable to capture complex patterns and relationships within the data (LeCun et al., 2015). Many algorithms have been proposed, demonstrating the great potential of ANNs in several forms; however, they also highlight significant challenges related to explainability (Silva et al., 2020).

The ANN architecture has evolved through the years leading to the development of more sophisticated models as Convolutional Neural Network (O'Shea & Nash, 2015) used mostly in computer vision or Recurrent Neural Networks primarily used for sequential data (Salehinejad et al., 2018). They have been applied in various road safety studies, including intersection-level risk modelling (Hu et al., 2020) and crash risk modelling at road segment (Wang et al., 2025). Moreover, sophisticated ANN architectures have been leveraged for real-time crash occurrence prediction (P. Li & Abdel-Aty, 2022) enabling dynamic and timely safety analysis.

This evolution has further led to the development of Graph Neural Networks (GNN) extending ANNs to graph-structured data, since many relationships in science and engineering can be represented as graphs (Scarselli et al., 2009). In transportation safety, most studies using GNNs aim to predict where and when crashes are likely to occur, treating transportation safety as a spatio-temporal forecasting problem (H. Li et al., 2024). Existing safety-related GNN approaches are centred on crash prediction using historical crash records (Nippani et al., 2024), while less attention has been given to fine-grained modeling of intersection-level or road-level safety using telematics indicators combined with roadway topology, despite telematics data increasingly being considered due to the scarcity of crash data and the relative ease of its collection (Joshi et al., 2025; Ziakopoulos et al., 2020). To address this gap, this study integrates graph-based representations with aggregated telematics features at the node level to generate node embeddings that capture both structural and behavioral characteristics of intersections. These embeddings preserve the road safety context from neighboring nodes and enable more effective downstream learning for tasks such as classification or clustering (Xu, 2021).

In road safety research, clustering techniques have been widely applied to analyze historical crash data, enabling the identification of distinct crash pattern groups at junctions (Nitsche et al., 2017) and the delineation of benchmark and high-risk regions to support targeted road safety interventions based on crash causation factors (Maji et al., 2018). Despite these advances, the clustering of spatial entities characterized by driving behavior remains relatively limited in the literature, as well as their integration with GNNs. For instance, (Huang et al., 2022) propose a traffic node importance evaluation framework that first constructs a length-weighted road network graph, then learns low-dimensional node representations using network embedding techniques, such as node2vec, and finally applies clustering to identify critical nodes based on vehicle flow characteristics. However, this work focuses on network importance rather than road safety.

The present explores the use of expressive representation learning techniques, such as GNNs, within a node embedding task, to analyze telematics data mapped onto a network and focus on node-based analysis performed weekly to examine temporal trends. GNNs were selected because they naturally model road networks as graphs and can learn spatial dependencies and node interactions directly from data, unlike traditional methods that rely on predefined spatial structures. By leveraging a GNN model further structural information can be incorporated into the representations of the nodes within the road network, called embeddings, leading to richer insights, in contrast to node2vec, which learns static embeddings based solely on random-walk–derived co-occurrence patterns. The study applies a clustering algorithm to the node embeddings to identify weekly partitions of node characteristics within a network and a comparison was made with a simpler clustering approach using raw node features alone, thereby demonstrating how embedding improves clustering performance and the analysis overall. Doing so, it is possible to identify improved and timely unsafe areas to provide a dynamic and proactive framework for road safety monitoring, supporting more effective interventions and resource allocation in traffic safety management.

Notably, the set of nodes characterized by telematics can vary from week to week, since diverse trip paths are contained within the weekly datasets. Despite this variability, the approach based on the GNN model consistently led to partitions with the same number of clusters each week. This framework provides a robust and reliable methodology for deriving spatial insights into driving behavior patterns, which can inform urban planning, traffic management, and road safety in real-world transportation networks, for example, helping prioritizing interventions at specific nodes or areas.

The paper is organized as follows: in *Data Collection and Preparation* the data sources are presented and the preprocessing steps detailed to establish a road network characterized by driving behavior data. *Methods* section discusses the methodologies employed to enhance the representation of network nodes and to identify a node-based partition of the network. In *Results and Discussion* the approaches are evaluated, their performances compared, and the findings interpreted to provide an overview of the results. Finally, the study is summarized and some takeaways are highlighted in the *Conclusions*.

2. Data Collection and Preparation

The present work is based on analyzing telematics data collected through a smartphone application developed by OSeven Telematics (OSeven, 2025), that records driver behavior using smartphone hardware sensors, offering a way for road safety analysis to avoid using crash data, enabling a proactive analysis rather than a reactive one. OSeven utilizes a variety of application programming interfaces (APIs) to retrieve sensor data, temporarily stored in the smartphone's database, and subsequently transmitted to the central back-end database (Kontaxi et al., 2023). The data are anonymized and stored in compliance with Greek and European data protection regulations (GDPR) and the APIs used support user authentication and encryption to prevent unauthorized access. Moreover, the driving behaviour data are pre-processed using advanced ML algorithms, which are proprietary and cannot be disclosed due to intellectual property restrictions (Papadimitriou et al., 2019).

The provided dataset consists of anonymized trip data from over 13,000 trips within a central area of the Athens metropolitan region, collected over the last four months of 2024 at a frequency of 1 Hz. The data provided contains trip coordinates along with per-second speed data, periodic binary flags (1 or 0) denoting the presence of harsh events, speeding or mobile use and it also includes the intensity of harsh events on a scale from 1 to 3. These metrics are generated by proprietary ML algorithms developed by OSeven, whose accuracy has been validated against OBD data, on-road tests and literature benchmarks. However, the dataset does not have an endless coverage, hence a coordinate bounding box can be defined in order to extract geometry features from OpenStreetMap (OSM) (OpenStreetMap, 2025), which is a free and editable global map created by volunteers and released under an open-content license.

The bounding box allows the extraction of a graph from OSM that includes node and edge features stored in two different datasets. A query was set to ensure that nodes are “true” edge endpoints (i.e., intersections or dead-ends) and the edges are the links between these nodes (Boeing, 2024).

Telematics were aggregated to the OSM entities by summing or averaging, depending on the feature’s nature; for instance, binary flags were summed to have a total count per entity, while continuous features (i.e., speed) were averaged. An enhanced buffer-based approach was used to aggregate telematics data at network nodes, where observations falling within a defined buffer area with a 50 meters radius were used to characterize each node, but only if they lay on the edges connected to the originating buffer node. For network edges, a closest-edge method was applied: each telematics observation was first matched to its nearest edge, and then data were aggregated per edge.

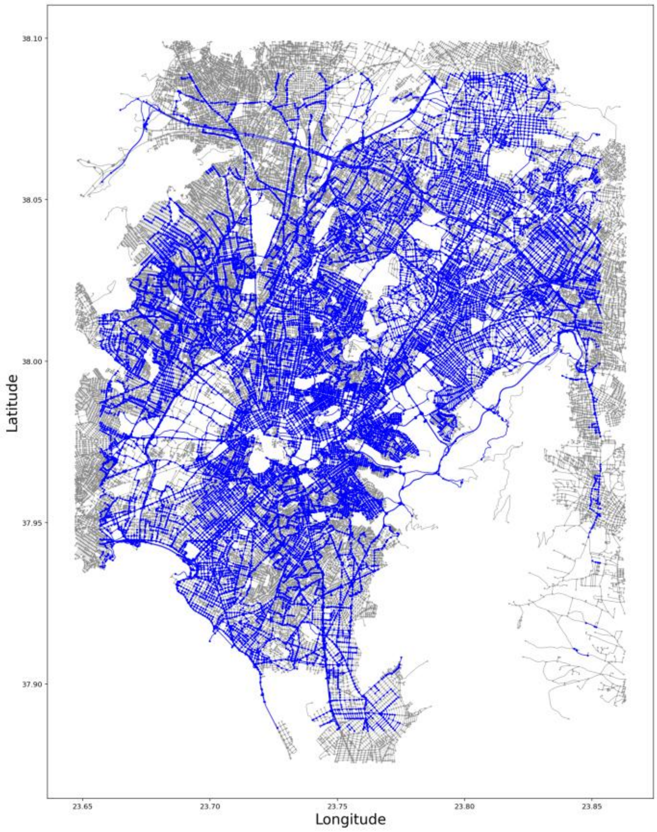

Fig. 1. Road Network with Telematics-Characterized Nodes and Edges.

Figure 1 illustrates the coverage of OSeven data, showing how telematics data coming from the trips contained in the provided dataset allowed for a characterization of certain edges and nodes within the selected study area.

The study area encompasses more than 40 municipalities, containing the metropolitan city of Athens. It covers approximately 400 km² and, according to OSM data, has a population of over 3,000,000 residents, representing around 30% of Greece's total population. The existing network of the metropolitan area of Athens, which incorporates the present study area, is car oriented. The layout allows the penetration of the compact urban core by primary arterials and urban motorways. As a result, the area presents unfavorable conditions for pedestrians, cyclists, and vulnerable social groups, since the intensive traffic flow passes through the intermunicipal and metropolitan centers as well as through numerous neighborhoods (Tsigdinos & Vlastos, 2021).

Prior to the spatial aggregation described above, the data were first aggregated by time window, enabling temporal aggregation and the creation of time-specific road networks characterized by telematics nodes and edges. For this study, a weekly time window was chosen, resulting in 18 weekly datasets covering the final 18 weeks of 2024. Segmenting the original dataset by week balances temporal resolution with data sparsity: daily aggregation could result in sparse networks, while monthly aggregation might obscure short-term variations in driving behavior. Additionally, the data covers only four months, which is insufficient to capture longer-term temporal trends or seasonal effects; therefore, a monthly analysis would be unreliable, whereas weekly aggregation allows for meaningful temporal patterns without excessive sparsity. At this stage, weekly graphs were selected including all edges characterized by telematics data, along with additional features derived from OSM. All nodes representing the initial and final nodes of these edges were also included.

3. Methods

This study aims to partition a road network by time windows (e.g., weekly) using telematics and geometric features. The methodologies employed include clustering analysis, silhouette analysis for clustering evaluation and the use of Graph Neural Networks (GNNs) to generate node embeddings. Each method is illustrated in detail below.

3.1. K-Means Clustering

The present study employs a clustering technique to find a partition of the presented data. Clustering is a type of unsupervised machine learning technique that groups objects into clusters, such that objects within each cluster are more similar to each other than to those in other clusters. One of the most commonly used clustering techniques is the K-Means algorithm, described in detail in (Steinley, 2006), that involves examining the similarity of points within a cluster relative to all other points in the same cluster (Steinbach et al., 2000). This algorithm was selected because of its simplicity, computational efficiency, and ability to produce well-defined, interpretable clusters. The K-means algorithm requires K, the number of clusters, as input, which is often a matter of domain knowledge or expert judgement. One method to determine the optimal K is to plot the silhouette score against several values of K and choose the maximum.

The silhouette score (Shahapure & Nicholas, 2020) is a commonly used evaluation metric for clustering, ranging from -1 to 1, and it is calculated for each data point in the dataset. A score near +1 means the data point is correctly clustered, near to 0 suggests potential overlap, and near -1 indicates wrong cluster. The average score provides information about whether the clusters are well-separated or overlapping. K-Means clustering was applied to the weekly node datasets to identify patterns and partition the node-based data into distinct groups, potentially revealing insights into high-risk and low-risk areas.

3.2. Node Embedding

In order to obtain a more informative representation of each node, capturing not only its individual features but also its topological relationships within the network and the connected edge features, a GNN was employed. A GNN is, in essence, an extension of existing ANNs suitable for data represented in graph domains (Scarselli et al., 2009). In this work, the GNN model was used to perform a node embedding task aiming to encode each node as a numeric vector, capturing its graph position and the structure of its local neighborhood context (Hamilton, 2020).

A GNN layer updates every node representation by aggregating the neighboring representations. Given an input set of node representations $\{h_i \in \mathbb{R}^d | i \in \mathcal{V}\}$ and the set of edges $\mathcal{E}$, the GNN layer outputs a new set of node representations $\{h_i{}' \in \mathbb{R}^{d'} | i \in \mathcal{V}\}$, through a combination of a function and an aggregation mechanism. This aggregation can involve methods such as averaging neighboring node features or applying attention to weight neighbors differently. The updated node representation is computed as:

$$h_i{}' = f_{\emptyset}\left(h_i, AGGREGATE\left(\{h_j \mid j \in \mathcal{N}_i\}\right)\right)$$

Where $\mathcal{N}_i = \{j \in \mathcal{V} | \ (j, i) \in \mathcal{E}\}$ are the neighbors of the node i.

One of the first models introduced in the GNN field was the Graph Convolution Networks (GCN) which uses a simple averaging aggregation combined with a convolution operation (Kipf & Welling, 2017). Improvements were made by leveraging the attention mechanisms leading to the Graph Attention Networks (GAT), which assigns different weights to neighboring nodes (Veličković et al., 2018). Further advancements resulted in the second version of the GAT model (GATv2), using what is known as dynamic attention, rather than static attention (Brody et al., 2022).

The GAT model (Veličković et al., 2018) works by applying a linear transformation, parametrized by a weight matrix $\boldsymbol{W}$, to every node. Then self-attention on the nodes are performed computing attention coefficients, indicating the importance of node j's features to node i. The coefficients are passed through a softmax function to normalize them across all choices of j.

The attention mechanism α is a single-layer feedforward ANN, parametrized by a weight vector $\bar{a}$, and applying the

LeakyReLU nonlinearity. Fully expanded, the coefficients computed by the attention mechanism are expressed as:

$$\alpha_{ij} = \frac{e^{(LeakyReLU(\bar{a}^T[\boldsymbol{W}\bar{h_i}||\boldsymbol{W}\bar{h_j}]))}}{\sum_{k\in\mathcal{N}_i} e^{(LeakyReLU(\bar{a}^T[\boldsymbol{W}\bar{h_i}||\boldsymbol{W}\bar{h_k}]))}}$$

Where || is the concatenation operation. The attention coefficients are then used to compute a linear combination of the features corresponding to them, to serve as the final output features for every node, after potentially applying an activation function σ to introduce nonlinearity:

$$h_i{}' = \sigma\left(\sum_{j\in\mathcal{N}_i} \alpha_{ij}\,\boldsymbol{W}\bar{h_j}\right)$$

Authors in (Veličković et al., 2018) found that extending the mechanism to employ multi-head attention was beneficial, hence the coefficients might be computed K times and the resulting node representations can then be either concatenated or averaged across the heads.

However, authors in (Brody et al., 2022) further improved this model, introducing a dynamic attention by simply modifying the order of internal operations in GAT and introducing GATv2, resulting in a more expressive attention mechanism. The attention function in the GATv2 model directly transforms the combined features together, allowing the model to learn a more flexible, dynamic attention function.

Fully expanded, the coefficients computed by the attention mechanism with GATv2 are expressed as follows:

$$\alpha_{ij} = \frac{e^{\bar{a}^T\cdot LeakyReLU(\boldsymbol{W}[\bar{h_i}||\bar{h_j}])}}{\sum_{k\in\mathcal{N}_i} e^{\bar{a}^T\cdot LeakyReLU(\boldsymbol{W}[\bar{h_i}||\bar{h_k}])}}$$

Moreover, when available, multi-dimensional edge features may be concatenated with node features before computing the attention coefficients, in order for them to affect the attention mechanism. The GATv2 model can be used to perform a node embedding task for each week and obtain weekly node representations involving edge features and network topology, besides the node raw features.

The model was trained in a self-supervised manner using a contrastive loss function specifically designed by the authors, inspired by prior work in self-supervised learning with contrastive loss function (Chen et al., 2020; Oord et al., 2019; Shen et al., 2023). The function works as follows:

1. For each node $i = 1$ to N.
    a. Calculate the cosine similarity with its neighbours $S_{ij}{}^+$ and with a randomly sampled set of non-neighbors $S_{ij}{}^-$.
    b. Scale these similarity scores using a temperature parameter and apply the exponential function to the scaled values to obtain $e^{\frac{S_{ij}{}^+}{\tau}}$ and $e^{\frac{S_{ij}{}^-}{\tau}}$.
    c. Sum the values over neighbours $\sum_{j\in P_i} e^{\frac{S_{ij}{}^+}{\tau}}$ and non-neighbours separately $\sum_{j\in N_i} e^{\frac{S_{ij}{}^-}{\tau}}$.
    d. Compute the ratio of neighbour sum to total sum, then apply logarithm:
    $$\mathcal{L}_i = -\log\left(\frac{\sum_{j\in P_i} e^{\frac{S_{ij}{}^+}{\tau}}}{\sum_{j\in P_i} e^{\frac{S_{ij}{}^+}{\tau}} + \sum_{j\in N_i} e^{\frac{S_{ij}{}^-}{\tau}}}\right).$$
2. Average over all nodes $\mathcal{L}_{TOT} = \frac{\sum_i \mathcal{L}_i}{N}$.

The function aims to optimize the learning process to obtain node embeddings that reflect shared characteristics and connectivity. Each weekly model was trained with 10 epochs using subgraphs sampled by PyG NeighborLoader and the optimization was performed with Adam optimizer, resulting in weekly node embeddings which will be used as input for K-Means.

4. Results and Discussion

After aggregating telematics data per node and per edge on a weekly basis, following the aforementioned steps, K-means clustering was applied to the node features, shown in Table 1. Roadway geometric features were derived from OSM, including the number of streets connected to each intersection. The remaining node features in Table 1 were calculated from aggregated telematics data. The total number of trips (Trips_count) was included as a separate feature to capture traffic demand and exposure levels at each location.

Table 1. Node Features.

Features	Description
Street_Count	Number of streets connected to the intersection
SmoothenedSpeed	Average speed of vehicles near the node
SpeedingFlag	Count of speeding events near the node
Mobile_usage	Number of instances of phone usage near the node
Harsh_acc	Number of harsh acceleration events near the node
Harsh_brk	Number of harsh braking events near the node
Event_intensity	Average intensity of events near the node
Trips_count	Number of trips recorded near the node

After scaling the features on a weekly basis in order to evaluate the separation of clusters relative to each week's distribution, a silhouette analysis was perfomed across the weekly raw datasets. Specifically, the silhouette score was computed for several values of K for the K-Means, across all 18 weeks. This approach allows us to compare how clusters are positioned within the weekly context, even when the underlying feature distributions vary substantially from week to week.

Table 2 presents the silhouette scores for each week. For most weeks, a natural partition involves only two natural groups. However, for some weeks, the highest silhouette score occurs with a large number of clusters and the score itself is very low, indicating that no clear natural grouping exists in the data for those periods.

Table 2. Silhouette Analysis on Weekly Raw Features.

Week	Optimal Clusters	Max Silhouette Score	Silhouette Score for k=2	Rank of k=2 in Silhouette Scores
2024-08-29	2	0.6116	-	-
2024-09-05	2	0.6079	-	-
2024-09-12	2	0.6042	-	-
2024-09-19	9	0.3905	0.2678	9
2024-09-26	10	0.3984	0.1890	9
2024-10-03	2	0.6264	-	-
2024-10-10	2	0.6101	-	-
2024-10-17	2	0.6227	-	-
2024-10-24	2	0.6137	-	-
2024-10-31	9	0.3940	0.2660	9
2024-11-07	2	0.6002	-	-
2024-11-14	3	0.5928	0.5871	2
2024-11-21	2	0.5758	-	-
2024-11-28	7	0.3715	0.2700	9
2024-12-05	2	0.6141	-	-
2024-12-12	2	0.5916	-	-

2024-12-19	2	0.6042	-	-
2024-12-26	2	0.6588	-	-

The value of K=2 was chosen for all weeks, disregarding those with a large number of clusters. Most of the cases where the optimal K was not 2 were characterized by a poor cluster quality, that is the silhouette score being lower than 0.4, along with an optimal K greater than 7 which would complicate the interpretation of the results. In one specific case, although the optimal K was 3, K=2 was selected as the silhouette score was nearly identical and offered a simpler clustering solution.

The comparison of the centroids across weeks is shown in Figure 2. To be consistent across the time windows, the smaller cluster was plotted in red, assuming that it would represent more unsafe driving behavior, since these patterns are expected to be rare and thus form less populated cluster. On the other hand, the larger cluster was plotted in green, assuming it would represent the one containing the majority of nodes where drivers tend to exhibit safer behavior.

Harsh acceleration, harsh braking events, average event intensity, and trip count exhibit a similar trend, showing negative peaks when clustering performance is poor (i.e., when the silhouette score drops below 0.50). This indicates that some features become less distinguishable between clusters under weak clustering conditions. For the remaining features, the figure shows abrupt changes in the centroid trends, particularly in terms of differences between centroids, when the clustering performances are poor, indicating that the features become unstable when the separation between clusters fluctuates over different time windows.

At this stage, the GATv2 model was used by employing multi-head attention mechanisms with three attention heads instead of a single attention head. The edge features are the same listed in Table 1, excluding street_count and supplemented by four additional features, either directly derived from OSM or encoded from OSM attributes:

- Edge length (directly from OSM geometry)
- Direction of movement (one or both directions; derived from OSM attributes and expressed as a binary column)
- Two binary features were created via one-hot encoding from a three-category road type variable (service, urban, or rural), manually encoded by inspecting the corresponding OSM attribute.

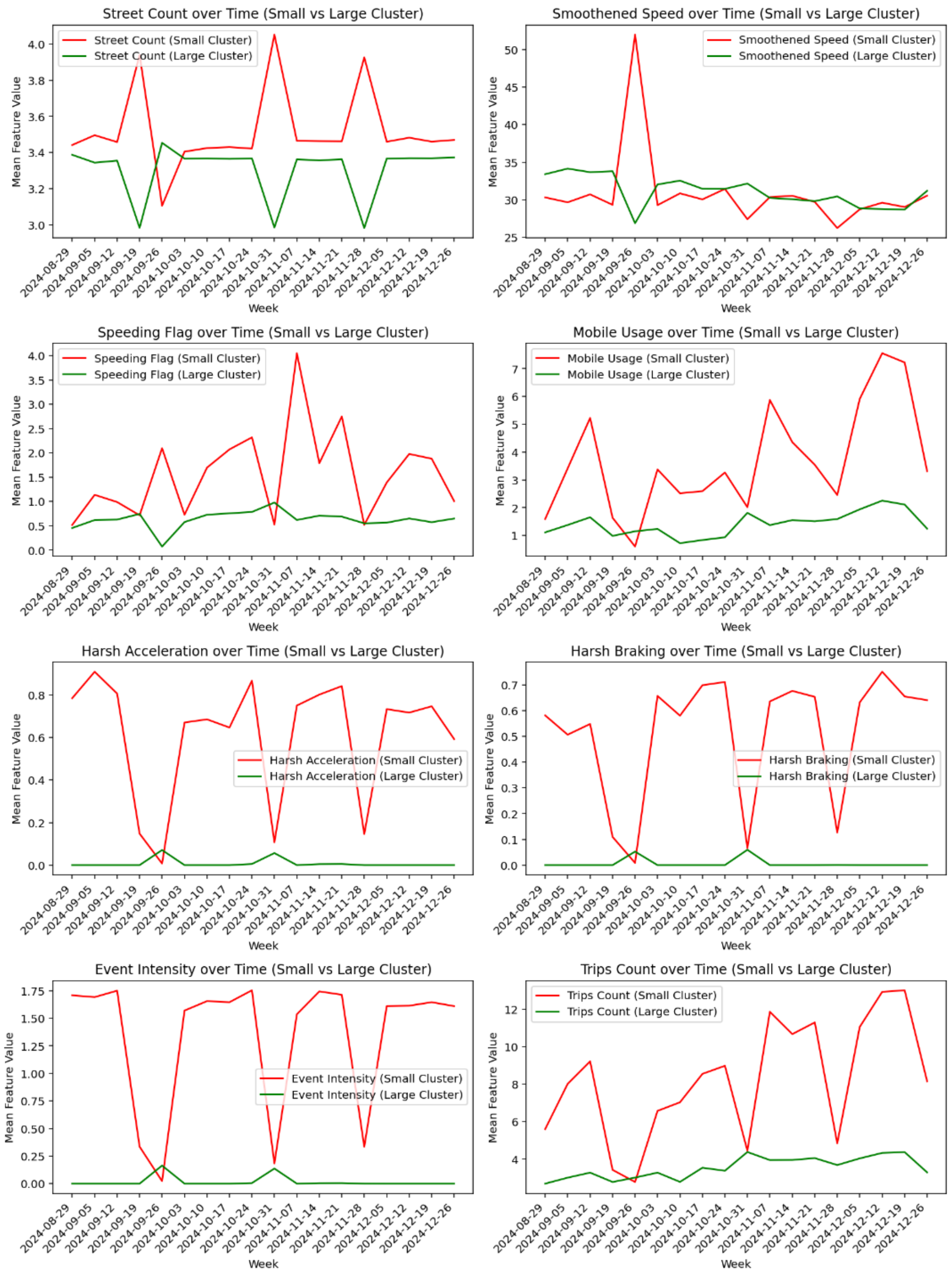

Fig. 2. Feature Trends of Centroids Over Time — Raw Feature Clusters.

A simple neural network was defined using two GATv2 layers, each followed by a LeakyReLU activation function, in accordance with the official related documentation of PyTorch library. The silhouette analysis previously illustrated was performed again on the node embeddings obtained by training the GATv2 model.

The number of clusters remains stable across the 18-week period, the GATv2 model is able to capture node representations that support consistent and natural groupings over time. K-Means applied to the node embeddings identified two distinct clusters each week. Additionally, the silhouette scores show a notable improvement, having a peak at 0.91 in one week and never falling below 0.60, still comparable to the best weekly performance achieved when clustering directly on raw features.

Table 3. Silhouette Analysis on Weekly Embeddings.

Week	Optimal Clusters	Max Silhouette Score
2024-08-29	2	0.6271
2024-09-05	2	0.6425
2024-09-12	2	0.9137
2024-09-19	2	0.8520
2024-09-26	2	0.6506
2024-10-03	2	0.6323
2024-10-10	2	0.7660
2024-10-17	2	0.7175
2024-10-24	2	0.6463
2024-10-31	2	0.6222
2024-11-07	2	0.6028
2024-11-14	2	0.6362
2024-11-21	2	0.6482
2024-11-28	2	0.7543
2024-12-05	2	0.6695
2024-12-12	2	0.6176
2024-12-19	2	0.6915
2024-12-26	2	0.6728

However, while clustering performance through node embeddings are improved, interpretability is lost due to the abstraction introduced by the embedding process. To bridge the gap between interpretability and performance, cluster labels were mapped back to the raw feature dataset and the features were averaged within each cluster across all weeks, resulting in weekly pseudo-cluster centroids, so called because these averages do not represent true centroids in the original feature space. These pseudo-cluster centroids were plotted as shown in Figure 3, to analyze temporal trends.

Following the same reasoning previously illustrated, the smaller cluster was plotted in red and the larger cluster in green. The larger cluster tends to be more stable over time and to show less variability with lower mean values for almost all features compared to the smaller cluster.

The larger cluster appears to comprise a more conservative group of drivers with lower speeds, almost zero speeding events, generally less mobile usage, and fewer harsh events over time. The smaller cluster represents a more dynamic group, characterized by more frequent speeding events, higher driving speeds overall, increased phone usage while driving, and a greater number and intensity of harsh driving events. The smaller cluster can be defined as a group of riskier areas characterized by hazardous behaviors, whereas the large cluster represents the majority of nodes with safer behavior. Moreover, these areas appear to be highly trafficked as well, which further increases risk exposure.

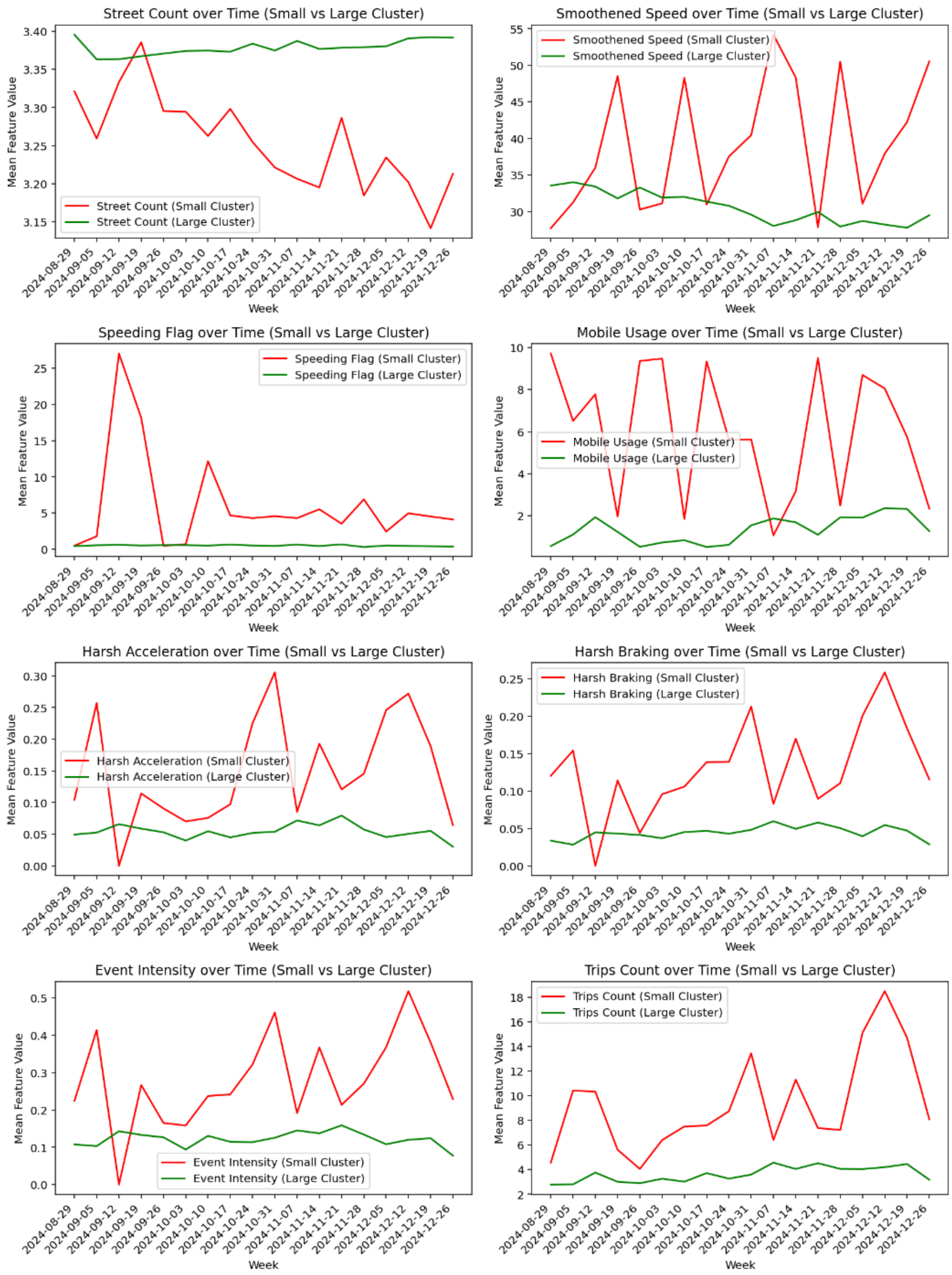

Fig. 3. Feature Trends of Centroids Over Time — Embedding Clusters.

When examining the temporal trends, it is worth noting that the week starting on September 12th, which is the one having the highest silhouette score among all weekly clusterings, exihibits a lower value of harsh events compared to the larger cluster. Despite this finding, it presents a significant peak in speeding events which may have driven the clustering outcome. Although the count of harsh events is very low, the elevated amount of speeding events hints towards the definition of this cluster as part of the hazardous ones, as the drivers tend to speed more frequently.

The larger group of nodes exhibits a more stable temporal pattern, indicating that driver behavior near these intersections remains relatively consistent across weeks. In contrast, the smaller clusters correspond to intersections where driving behavior varies across weeks. This result appears to be linked to the influence of different hazardous behaviors on group-level risk over time. For instance, during the week of November 28th, both clusters exhibited comparable levels of mobile phone use, but drivers in the smaller one were characterized by a higher speed. In the following week of December 5th both groups showed a similar number of speeding events and comparable average speeds, while the smaller cluster displayed higher levels of mobile phone use.

This modelling approach can be seen as a diagnostic tool, which enables a shift from reactive crash reporting to proactive identification of unsafe "red" clusters. The larger cluster, showing stable behaviour over time, suggests lower priority for intervention, whereas the smaller cluster exhibits stronger temporal variability. In this red cluster, risk may be seen as alternately associated with harsh events, speeding, or mobile phone use, suggesting different hypothetical mitigation strategies depending on the dominant behaviour and the time window.

Overall, the K-Means clustering based on GNN-generated node embeddings consistently outperforms the K-Means clustering on raw features across all weekly partitions, as reflected by higher silhouette scores for every week in the dataset, enhancing the natural gruping for each time window. Furthermore, when clustering on raw features fails to identify a partition in the network due to very low silhouette scores, the GNN-based approach finds clusters with reasonable performance. Thus, leveraging node embeddings not only improves clustering quality but also finds groups that are difficult to identify using raw features alone.

Additionally, this approach leads to a more stable number of clusters over time indeed, with the silhouette scores indicating a fixed number of 2 clusters for each week within the analyzed period.

By employing GNNs for node embedding, it is possible to improve the node representation, incorporating an additional level of information brought from the neighborhood context, and utilizing this improved node representation for more accurate and meaningful partitions of the network, consistent across all time windows analyzed.

This richer representation enables the clustering algorithm to better distinguish node-based patterns within the road network, while involving spatial relantionships among nodes. Indeed, while working with raw features leads to considering nodes in isolation, the GNN-based embeddings capture spatial relationships through the neighborhood context.

This approach shows great promise for proactive spatial risk analysis, where understanding nuanced spatial relationships is essential for effective decision-making. Spatial road safety analyses can leverage telematics data on driving behavior to examine driver characteristics across locations, helping to identify areas where drivers are prone to aggressive or careless actions. Graph-based deep learning methods improve the understanding of the network leading to a better partition in which nodes are grouped more effectively, while the segmentation by week provides insight into the different hazardous events that characterize data separation across different weeks.

Together, these refinements yield an improved partition that can inform targeted interventions, such as awareness campaigns or enforcement measures, to enhance road safety. In practice, more precise identification of unsafe locations allows policymakers to design effective interventions, improving the management of limited resources by ensuring they are allocated more efficiently, since previous research has shown that identifying and targeting crash hotspots is a key step in road safety management (Turki et al., 2022).

The results of this study might be generalized to other locations, since OSM offers globally consistent road infrastructure features. However, the applicability of the framework depends on the availability of aggregated telematics indicators, or alternatively, on adapting the whole pipeline to the available dataset. Therefore, while the methodology is transferable, the framework may need to be re-specified or adapted when applied to new datasets with different telematics features.

4. Conclusions

The current work proposes a novel approach for understanding the nature of graph structured data. While several simple and more composite clustering algorithms have been explored within the road safety field, this research demonstrates the advantages of exploiting GNNs for node embedding, rather than relying on shallow embedding methods such as node2vec, to obtain richer node representations that are subsequently used as input to clustering algorithms. Additionally, the network partition was performed on a weekly basis to explore the temporal trends of the cluster characteristics.

A simple clustering algorithm, K-Means, was used to compare the clustering results on raw features with those obtained on node embeddings. The methodology demonstrates how GNN-generated node representations improve clustering performance across all weeks, while also stabilizing the number of optimal clusters within the data over time.

Clustering node embeddings effectively captures the separation of spatial entities within a telematics-informed road network, leading to a better identification of locations where drivers tend to behave more aggressively or carelessly. These areas can then be targeted for interventions to help reduce hazardous driving behaviors and enhance overall traffic safety.

The study is grounded in the use of K-Means clustering, the GATv2 layer and a weekly temporal aggregation window. Further research could benefit from exploring alternative clustering techniques and a broader range of GNN architectures. Given the diversity of methods available for graph representation learning, such extensions may yield deeper insights and potentially improve model performance or interpretability. In addition, examining different temporal resolutions beyond the weekly window might help to capture patterns occurring at different time scales.

Additionally, the 50-meter radius for the buffer could be further evaluated, and alternative values tested across different spatial scales to assess sensitivity. Ultimately, the feature Trips_count may bias the clustering toward traffic volume effects, as clusters with higher risk-related events also show higher traffic volume, suggesting that the results may reflect traffic density rather than behavioral risk. Future studies should therefore consider normalizing event features by Trips_count to better isolate underlying behavioral patterns.

Acknowledgements

This research is based on work carried out within the IVORY project. The project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement No 101119590.

References

Boeing, G. (2024). Graph Simplification Solutions to the Street Intersection Miscount Problem (arXiv:2407.00258). arXiv. https://doi.org/10.48550/arXiv.2407.00258

Brody, S., Alon, U., & Yahav, E. (2022). How Attentive are Graph Attention Networks? (arXiv:2105.14491). arXiv. https://doi.org/10.48550/arXiv.2105.14491

Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A Simple Framework for Contrastive Learning of Visual Representations (arXiv:2002.05709). arXiv. https://doi.org/10.48550/arXiv.2002.05709

Hamilton, W. L. (2020). Graph Representation Learning. Morgan & Claypool Publishers.

Hu, J., Huang, M.-C., & Yu, X. (2020). Efficient mapping of crash risk at intersections with connected vehicle data and deep learning models. Accident Analysis & Prevention, 144, 105665. https://doi.org/10.1016/j.aap.2020.105665

Huang, X., Chen, J., Cai, M., Wang, W., & Hu, X. (2022). Traffic Node Importance Evaluation Based on Clustering in Represented Transportation Networks. IEEE Transactions on Intelligent Transportation Systems, 23(9), 16622–16631. https://doi.org/10.1109/TITS.2022.3163756

Joshi, M., Bamney, A., Wang, K., Zhao, S., Ivan, J., & Jackson, E. (2025). Analyzing the Suitability of Vehicle Telematics Data as a Surrogate Safety Measure for Short-Term Crashes. Transportation Research Record, 2679(2), 489–504. https://doi.org/10.1177/03611981241263341

Kipf, T. N., & Welling, M. (2017). Semi-Supervised Classification with Graph Convolutional Networks (arXiv:1609.02907). arXiv. https://doi.org/10.48550/arXiv.1609.02907

Kontaxi, A., Tzoutzoulis, D.-M., Ziakopoulos, A., & Yannis, G. (2023). Exploring speeding behavior using naturalistic car driving data from smartphones. Journal of Traffic and Transportation Engineering (English Edition), 10(6), 1162–1173. https://doi.org/10.1016/j.jtte.2023.07.007

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Li, H., Zhao, Y., Mao, Z., Qin, Y., Xiao, Z., Feng, J., Gu, Y., Ju, W., Luo, X., & Zhang, M. (2024). Graph Neural Networks in Intelligent Transportation Systems: Advances, Applications and Trends (arXiv:2401.00713). arXiv. https://doi.org/10.48550/arXiv.2401.00713

Li, P., & Abdel-Aty, M. (2022). Real-Time Crash Likelihood Prediction Using Temporal Attention–Based Deep Learning and Trajectory Fusion. Journal of Transportation Engineering, Part A: Systems, 148(7), 04022043. https://doi.org/10.1061/JTEPBS.0000697

Maji, A., Velaga, N. R., & Urie, Y. (2018). Hierarchical clustering analysis framework of mutually exclusive crash causation parameters for regional road safety strategies. International Journal of Injury Control and Safety Promotion, 25(3), 257–271. https://doi.org/10.1080/17457300.2017.1416485

Nippani, A., Li, D., Ju, H., Koutsopoulos, H. N., & Zhang, H. R. (2024). Graph Neural Networks for Road Safety Modeling: Datasets and Evaluations for Accident Analysis (arXiv:2311.00164). arXiv. https://doi.org/10.48550/arXiv.2311.00164

Nitsche, P., Thomas, P., Stuetz, R., & Welsh, R. (2017). Pre-crash scenarios at road junctions: A clustering method for car crash data. Accident Analysis & Prevention, 107, 137–151. https://doi.org/10.1016/j.aap.2017.07.011

Oord, A. van den, Li, Y., & Vinyals, O. (2019). Representation Learning with Contrastive Predictive Coding (arXiv:1807.03748). arXiv. https://doi.org/10.48550/arXiv.1807.03748

OpenStreetMap. (2025). About OpenStreetMap—OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/About_OpenStreetMap

OSeven. (2025). Oseven.io. Oseven.Io. https://oseven.io/

O'Shea, K., & Nash, R. (2015). An Introduction to Convolutional Neural Networks (arXiv:1511.08458). arXiv. https://doi.org/10.48550/arXiv.1511.08458

Papadimitriou, E., Argyropoulou, A., Tselentis, D. I., & Yannis, G. (2019). Analysis of driver behaviour through smartphone data: The case of mobile phone use while driving. Safety Science, 119, 91–97. https://doi.org/10.1016/j.ssci.2019.05.059

Salehinejad, H., Sankar, S., Barfett, J., Colak, E., & Valaee, S. (2018). Recent Advances in Recurrent Neural Networks (arXiv:1801.01078). arXiv. https://doi.org/10.48550/arXiv.1801.01078

Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., & Monfardini, G. (2009). The Graph Neural Network Model. IEEE Transactions on Neural Networks, 20(1), 61–80. IEEE Transactions on Neural Networks. https://doi.org/10.1109/TNN.2008.2005605

Shahapure, K. R., & Nicholas, C. (2020). Cluster Quality Analysis Using Silhouette Score. 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA), 747–748. https://doi.org/10.1109/DSAA49011.2020.00096

Shen, X., Sun, D., Pan, S., Zhou, X., & Yang, L. T. (2023). Neighbor Contrastive Learning on Learnable Graph Augmentation (arXiv:2301.01404). arXiv. https://doi.org/10.48550/arXiv.2301.01404

Silva, P. B., Andrade, M., & Ferreira, S. (2020). Machine learning applied to road safety modeling: A systematic literature review. Journal of Traffic and Transportation Engineering (English Edition), 7(6), 775–790. https://doi.org/10.1016/j.jtte.2020.07.004

Steinbach, M., Karypis, G., & Kumar, V. (2000). A Comparison of Document Clustering Techniques.

Steinley, Douglas. (2006). K-means clustering: A half-century synthesis. British Journal of Mathematical and Statistical Psychology, 59(1), 1–34. https://doi.org/10.1348/000711005X48266

Tsigdinos, S., & Vlastos, T. (2021). Exploring ways to determine an alternative strategic road network in a metropolitan city: A multi-criteria analysis approach. IATSS Research, 45(1), 102–115. https://doi.org/10.1016/j.iatssr.2020.06.002

Turki, Z., Ghédira, A., Ouni, F., & Kahloul, A. (2022). Spatio-temporal analysis of road traffic accidents in Tunisia. 2022 14th International Colloquium of Logistics and Supply Chain Management (LOGISTIQUA), 1–7. https://doi.org/10.1109/LOGISTIQUA55056.2022.9938067

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., & Bengio, Y. (2018). Graph Attention Networks (arXiv:1710.10903). arXiv. https://doi.org/10.48550/arXiv.1710.10903

Wang, Y., Wang, X., Wang, T., & Quddus, M. (2025). Freeway real-time crash prediction using floating car data. Transportation Research Part C: Emerging Technologies, 171, 105009. https://doi.org/10.1016/j.trc.2025.105009

WHO. (2023). Global status report on road safety 2023. https://www.who.int/publications/i/item/9789240086517

Xu, M. (2021). Understanding Graph Embedding Methods and Their Applications. SIAM Review, 63(4), 825–853. https://doi.org/10.1137/20M1386062

Ziakopoulos, A., Tselentis, D., Kontaxi, A., & Yannis, G. (2020). A critical overview of driver recording tools. Journal of Safety Research, 72, 203–212. https://doi.org/10.1016/j.jsr.2019.12.021

Ziakopoulos, A., & Yannis, G. (2020). A review of spatial approaches in road safety. Accident Analysis & Prevention, 135, 105323. https://doi.org/10.1016/j.aap.2019.105323