Security of Infocommunication Systems and Internet of Things

2026 Vol 4, No 1

https://doi.org/10.31861/sisiot2026.1.01013

Received 28 April 2026; revised 05 June 2026; accepted 19 June 2026; published 30 June 2026

Integrated Risk Assessment of Road Traffic Accident Clusters Based on Weighted Casualty Metrics

Bohdan Pashkovskyi*

Computer Systems and Networks Department, Ivano-Frankivsk National Technical University of Oil and Gas, Ivano-Frankivsk, Ukraine

*Corresponding author (E-mail: bohdan.pashkovskyi@nung.edu.ua)

ABSTRACT The paper addresses the problem of identifying and prioritizing hazardous road traffic accident concentration areas based on accident frequency and casualty severity. The relevance of the study is motivated by the increasing number of accidents with injuries and fatalities, as well as the limitations of approaches based solely on quantitative indicators without considering the severity of consequences. Under such conditions, treating accidents of different severity levels equally leads to distorted assessments of road safety and reduces the effectiveness of decision-making in traffic safety management. The study proposes an integrated risk assessment method based on weighted coefficients assigned to different categories of casualties. Minor injuries are assigned 1-point, severe injuries – 3 points, and fatalities – 9 points. Based on these coefficients, a generalized risk indicator is formed for both individual accidents and their aggregation within spatial clusters. The proposed approach is integrated into the process of clustering road traffic accidents using geographic coordinates and distance-based metrics. This allows not only the identification of accident concentration areas but also their ranking according to the level of danger. As a result, it becomes possible to detect critical locations, including those with a relatively low number of accidents but high severity of consequences. The practical implementation of the method is carried out within an information system for accident analysis, providing automated data processing, clustering, and report generation. The proposed method enables accident clusters to be ranked using both accident frequency and the severity-weighted consequences of the recorded accidents. The method was evaluated using accident records from 2017 to 2025 for the Ivano-Frankivsk urban territorial community, with a spatial threshold of 50m and a minimum cluster size of three accidents. The results show that casualty-weighted ranking differs from ranking based solely on accident frequency: clusters with fewer accidents but more fatalities and severe injuries may receive a higher priority. From the perspective of infocommunication systems, the proposed method can be used as an analytical component of an integrated road safety monitoring platform that combines distributed data acquisition, communication, centralized processing, and geospatial visualization. The system architecture also provides a basis for further integration with IoT-enabled traffic and road infrastructure devices.

KEYWORDS road traffic accident, accident cluster, integrated risk assessment, clustering, road safety.

I. INTRODUCTION

According to statistical data provided by the Department of Patrol Police of Ukraine, recent years have been characterized by a steady increase in the number of road traffic accidents involving injuries and fatalities. A significant proportion of these incidents involves vulnerable categories of road users, particularly pedestrians and cyclists. Under such conditions, the timely identification of accident concentration zones and hazardous road segments becomes a task of critical importance.

Existing methodologies for identifying such locations are predominantly based on counting the number of accidents within a specified time period and assessing their spatial concentration. However, this approach does not account for the severity of accident outcomes, which may lead to suboptimal prioritization in the planning and implementation of road safety measures.

Therefore, there is a clear need to develop an approach that considers not only the frequency of road traffic accidents but also their severity in terms of consequences for the life and health of road users.

From the perspective of infocommunication systems, road safety analysis requires the integration of heterogeneous data acquisition, communication, storage, processing, and visualization components. Accident-related information may be obtained from official reporting systems, mobile applications, intelligent traffic infrastructure, connected cameras, traffic counters, weather stations, and other Internet of Things devices. These distributed data sources can provide complementary information on traffic intensity, environmental conditions, road infrastructure status, and the circumstances of accident occurrence.

Accordingly, the proposed accident clustering and hazard assessment method can be considered an analytical component of an infocommunication road safety monitoring system. Such a system provides communication between distributed data sources and a centralized processing platform, where the collected data are validated, geocoded, stored, clustered, and analyzed. The integration of IoT-generated data can further improve the timeliness and contextual completeness of hazardous-location identification.

II. ANALYSIS OF RECENT RESEARCH AND PUBLICATIONS

The identification of road traffic accident concentration areas is commonly addressed using geographic information systems, spatial statistics, clustering algorithms, and predictive data analysis. A comprehensive review of spatial approaches in road safety demonstrated that crash analysis must account for spatial dependence, spatial heterogeneity, the selection of appropriate spatial units, and the structure of the road network [1].

Early GIS-based studies established the methodological basis for spatial accident analysis. Xie and Yan proposed Network Kernel Density Estimation, which calculates accident density directly within road-network space rather than over an unrestricted planar surface [2]. Anderson combined Kernel Density Estimation with K-means clustering to identify and classify road accident hotspots according to their spatial and environmental characteristics [3]. Erdogan et al. developed a GIS-supported accident analysis system that integrated accident records with spatial and attribute data [4]. These studies remain methodologically important, although their primary emphasis was on accident frequency and spatial density.

More recent research has focused on comparing and validating different hotspot identification techniques. Zahran et al. evaluated several GIS-based spatial analysis methods and compared the identified hotspots with the results of a road safety audit [5]. Their findings showed that hotspot identification depends considerably on the selected analytical method and spatial configuration. Density-based clustering methods have also received increased attention because they can identify irregularly shaped clusters and distinguish isolated observations from concentration areas. Islam et al. compared clustering approaches for road traffic crash data and recommended DBSCAN and OPTICS for datasets characterized by spatial noise and non-uniform cluster structures [6].

A separate research direction concerns the incorporation of crash severity into spatial hotspot analysis. Alam and Tabassum combined Getis–Ord Gi*, Moran’s I, and a crash severity index to identify and rank road traffic crash hotspots [7]. Their study demonstrated that crash frequency alone does not fully represent the danger associated with a location and that severity-related indicators can change the prioritization of hotspots. Gedamu et al. examined spatial autocorrelation separately for different crash severity categories using Global and Local Moran’s I [8]. Their results confirmed that severe and less severe crashes may form different spatial patterns.

Mesic et al. applied a weighted composite injury severity measure that included crashes, minor injuries, severe injuries, and fatalities and combined it with emerging hotspot analysis [9]. This approach made it possible to identify road sections where injury severity was persistent or increasing over time. However, the severity indicator was used mainly within a spatiotemporal statistical analysis, whereas the present study integrates casualty weights directly into the score calculated for every accident and subsequently aggregates these scores at the cluster level.

Temporal variation is another important component of contemporary crash analysis. Mohammed et al. used a space–time cube, emerging hotspot analysis, spatial autocorrelation, and geographically weighted regression to examine the spatial and temporal evolution of road traffic crashes [10]. Such methods distinguish persistent accident concentration areas from short-term or emerging patterns but generally require extended observation periods and more complex spatial data structures.

Machine-learning methods extend road safety analysis by incorporating road, traffic, and environmental characteristics. Berhanu et al. combined Random Forest prediction, crash-rate estimation, and spatial network analysis to identify accident-prone road segments and recommend safer routes [11]. These models can capture nonlinear relationships between accident occurrence and explanatory factors, although their results depend on the availability, completeness, and representativeness of training data.

The development of intelligent transport and infocommunication systems has also created new sources of road safety data. Tahir et al. proposed an event-driven traffic monitoring system that uses CCTV video analytics to detect and classify road traffic events and transmit relevant information for further processing [12]. Such systems demonstrate how distributed cameras, roadside devices, traffic counters, weather stations, and other Internet of Things components can supply data to centralized analytical platforms. A recent systematic review of GIS-based road traffic accident analysis likewise identified the integration of spatial analysis, heterogeneous datasets, predictive methods, and real-time monitoring as important directions for further research [13].

Previous studies by the author and co-author addressed the successive stages of automated accident data processing, including geocoding, data storage, information-system development, distance calculation, and spatial clustering [14-16]. Study [16] introduced an automated clustering procedure based on geographic coordinates and the Haversine distance. However, cluster prioritization in those studies was based primarily on the number and spatial concentration of recorded accidents.

The reviewed literature shows that spatial hotspot detection, severity assessment, temporal analysis, and intelligent monitoring are frequently implemented as separate analytical tasks. Some recent studies incorporate severity indices into GIS-based analysis [7-9], but there remains a need for a transparent method that assigns a hazard score to an individual accident and then aggregates this score within the spatial clusters produced by an automated information system.

The present study addresses this problem by introducing casualty-based weighting coefficients for minor injuries, severe injuries, and fatalities. The resulting integral score is calculated at both the accident and cluster levels and is used together with spatial clustering to rank accident concentration areas. The method is implemented as an analytical component of an infocommunication road safety system that can subsequently be extended through the integration of IoT-enabled traffic and infrastructure data sources.

III. METHODOLOGY

This study addresses the problem of automated identification of locations and segments with a high concentration of road traffic accidents based on their spatial characteristics and consequences. The input data consist of accident records obtained from official reporting forms, including the date and time of the incident, its type, causes, characteristics of the road environment, as well as the number of casualties categorized into minor injuries, severe injuries, and fatalities.

A. Input data and preprocessing. The input dataset is represented as a finite set of road traffic accident (RTA) records. Each record contains the accident location, casualty outcomes, temporal information, and descriptive attributes required for spatial analysis:

$$D = \{A_1, A_2, \ldots, A_n\}. \quad (1)$$

Each accident record A_i is represented as

$$A_i = (\varphi_i, \lambda_i, m_i, s_i, f_i, x_i), \quad (2)$$

where φ_i and λ_i are the latitude and longitude of the accident location; m_i, s_i, and f_i are the numbers of minor injuries, severe injuries, and fatalities, respectively; and x_i denotes additional attributes such as the date and time, accident type, street address, road-user category, and road-environment characteristics.

Before the analytical procedure is performed, the records are checked for missing or inconsistent casualty values and invalid coordinates. Textual accident locations are geocoded to obtain geographic coordinates. Records for which reliable coordinates cannot be obtained are excluded from spatial clustering or marked for manual verification. Latitude and longitude values are converted to radians before distance calculations.

B. Casualty-weighted hazard assessment. For every accident A_i, an individual hazard score h_i is calculated as a weighted sum of the casualty outcomes:

$$h_i = w_m m_i + w_s s_i + w_f f_i, \quad (3)$$

where w_m, w_s, and w_f are the weighting coefficients assigned to minor injuries, severe injuries, and fatalities. In the present study, these coefficients are defined as

$$w_m = 1, \quad w_s = 3, \quad w_f = 9. \quad (4)$$

Consequently, the accident-level hazard score is

$$h_i = m_i + 3s_i + 9f_i. \quad (5)$$

The coefficients define a transparent ordinal severity scale: a severe injury is assigned three times the weight of a minor injury, while a fatality is assigned three times the weight of a severe injury and nine times the weight of a minor injury. The score therefore does not estimate the monetary cost of an accident; it is used as a relative prioritization indicator.

For an identified accident cluster C_k, the aggregated casualty indicators are

$$M_k = \sum(A_i \in C_k)\, m_i, \quad (6)$$

$$S_k = \sum(A_i \in C_k)\, s_i, \quad (7)$$

$$F_k = \sum(A_i \in C_k)\, f_i. \quad (8)$$

The integral hazard score of cluster C_k is defined as

$$H_k = \sum(A_i \in C_k)\, h_i. \quad (9)$$

Using the selected casualty weights, this expression is equivalent to

$$H_k = M_k + 3S_k + 9F_k. \quad (10)$$

The number of accidents contained in the cluster is

$$N_k = |C_k|. \quad (11)$$

Thus, each cluster is described by the tuple $C_k = (N_k, M_k, S_k, F_k, H_k)$. The indicator N_k represents accident frequency, whereas H_k represents severity-weighted casualty consequences. A cluster with fewer accidents may therefore receive a higher priority when its accidents have resulted in more severe outcomes.

C. Spatial clustering method. The spatial relationship between accidents A_i and A_j is determined from the geographic distance between their coordinates. The distance is calculated using the Haversine formula, which accounts for the curvature of the Earth.

$$\Delta\varphi = \varphi_j - \varphi_i, \quad \Delta\lambda = \lambda_j - \lambda_i, \quad (12)$$

$$a_{ij} = \sin^2(\Delta\varphi/2) + \cos(\varphi_i)\cos(\varphi_j)\sin^2(\Delta\lambda/2), \quad (13)$$

$$d_{ij} = 2R \operatorname{atan} 2(\sqrt{a_{ij}}, \sqrt{(1 - a_{ij})}), \quad (14)$$

where R is the mean radius of the Earth and d_{ij} is the great-circle distance between the two accident locations. The clustering procedure is represented as an undirected proximity graph $G = (V, E)$, where every accident record corresponds to one vertex:

$$V = \{A_1, A_2, \ldots, A_n\}. \quad (15)$$

An edge is created between accidents A_i and A_j when the calculated distance does not exceed the predefined spatial threshold ε:

$$(A_i, A_j) \in E \iff d_{ij} \le \varepsilon. \quad (16)$$

In the experimental analysis, the spatial threshold was set to ε=50m. This value was selected based on an empirical comparison of several threshold values.

Equivalently, the adjacency relation can be expressed as

$$g_{ij} = \{\, 1, \text{ if } d_{ij} \le \varepsilon;\ \ 0, \text{ if } d_{ij} > \varepsilon \,\}. \quad (17)$$

An accident cluster C_k is defined as a connected component of graph G. Therefore, two accidents may belong to the same cluster either because they are directly within the distance threshold or because they are connected through a sequence of neighboring accident locations. This definition formalizes the iterative merging procedure used in the software implementation.

D. Algorithm of the proposed method. Algorithm 1. Spatial clustering and casualty-weighted hazard assessment.

Input: accident dataset D; spatial threshold ε; weights w_m, w_s, and w_f.

Output: clusters ranked by the integral hazard score H_k.

1. Validate the input records and obtain reliable geographic coordinates.
2. For each accident A_i, compute $h_i = m_i + 3s_i + 9f_i$.
3. Initialize an undirected graph G with one vertex for each accident.
4. For every unordered pair (A_i, A_j), calculate d_{ij} using the Haversine formula.

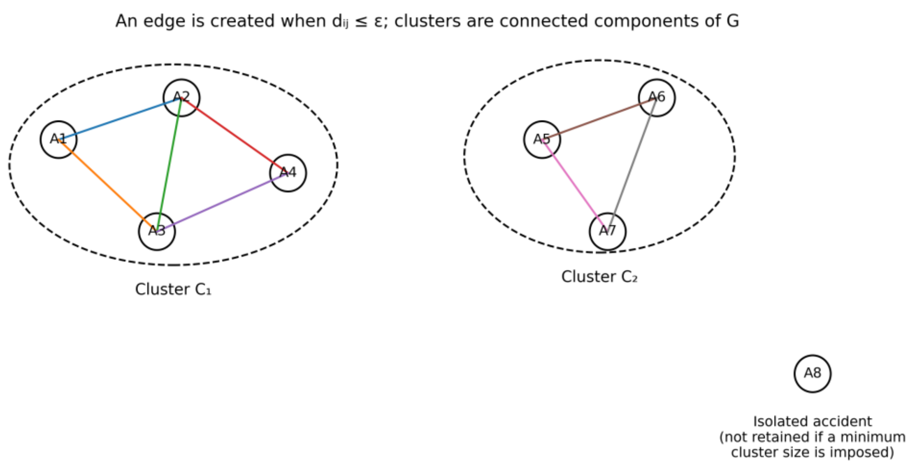

FIG. 1. Graph-based interpretation of spatial accident clustering. Edges are created when $d_{ij} \leq \varepsilon$, and clusters are the connected components of G.

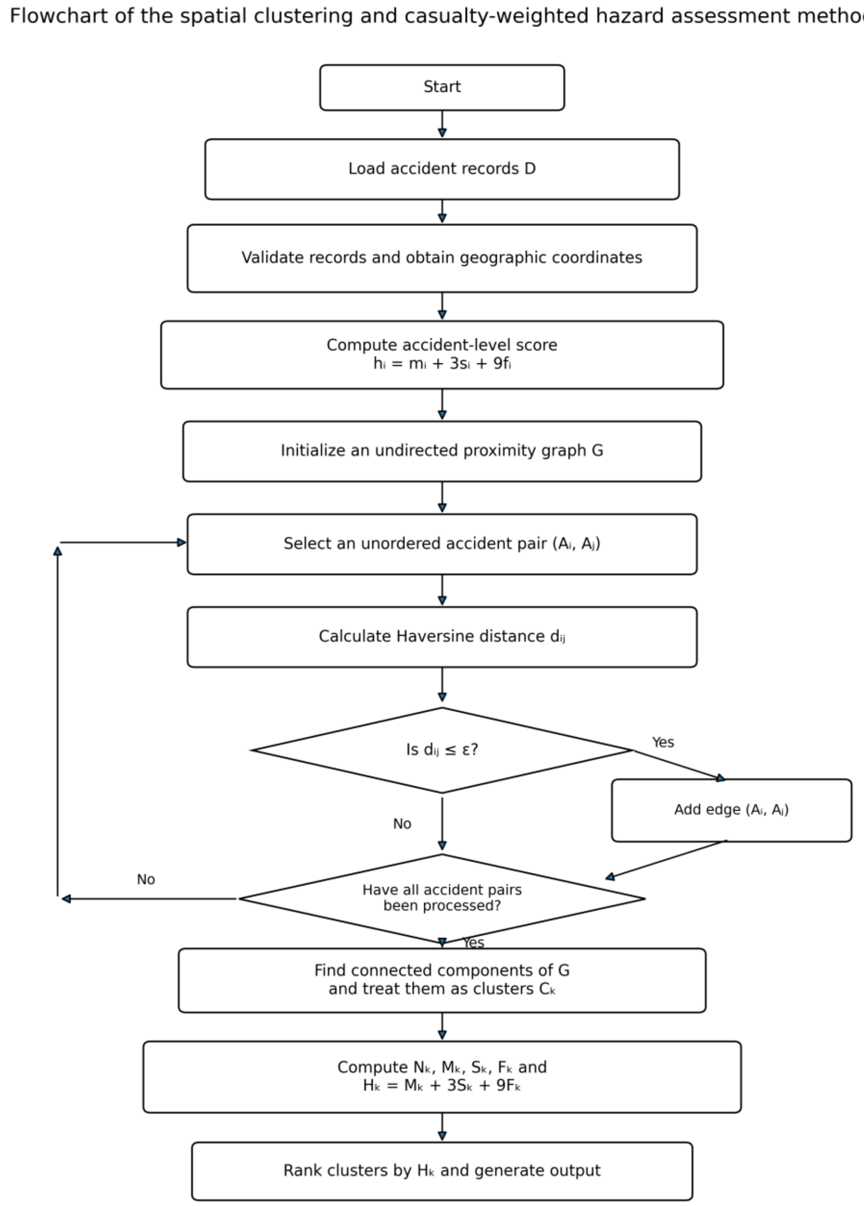

FIG. 2. Flowchart of the spatial clustering and casualty-weighted hazard assessment method.

5. If $d_{ij} \leq \varepsilon$, add the edge (A_i, A_j) to G.
6. Determine the connected components of G and treat each component as a cluster C_k.
7. For each C_k, calculate N_k, M_k, S_k, F_k, and $H_k = M_k + 3S_k + 9F_k$.
8. Rank the clusters in descending order of H_k.
9. Generate the analytical table, map visualization, and cluster report.

E. Computational complexity. The straightforward implementation evaluates every unordered pair of accident records. The number of pairwise distance calculations is

$$n(n - 1) / 2 \tag{18}$$

and the corresponding distance-comparison stage has a computational complexity of $O(n^2)$, where n is the number of accident records. Calculation of the individual accident scores and aggregation of cluster indicators require $O(n)$ operations. Consequently, the overall complexity of the implemented procedure is dominated by the pairwise spatial comparison and is $O(n^2)$.

For substantially larger regional or national datasets, the number of distance calculations may be reduced using spatial indexing, grid-based partitioning, an R-tree, or another neighborhood-search structure. These optimizations are proposed as future improvements and were not included in the implementation evaluated in this study.

F. Information system implementation and output. The proposed method was implemented as part of a web-based information system for road traffic accident analysis. The system stores accident records in a relational database, supports filtering by temporal, spatial, and descriptive attributes, calculates Haversine distances, constructs accident clusters, and evaluates the accident-level and cluster-level hazard scores.

For each cluster, the system produces the indicators N_k, M_k, S_k, F_k, and H_k. The clusters are ranked in descending order of H_k and presented through an analytical table and a map visualization. The simultaneous presentation of N_k and H_k enables the user to distinguish between locations characterized by frequent accidents and locations characterized by severe casualty outcomes.

The current implementation processes accident records obtained from official reporting sources. Integration with IoT-enabled traffic counters, connected road cameras, weather stations, and other distributed road-monitoring devices is treated as an architectural extension and a direction for further research rather than as functionality evaluated in the present study.

IV. RESULTS AND DISCUSSION

A. Experimental dataset and parameter settings. The proposed method was evaluated using road traffic accident records processed by the developed web-based information system. The report output used for this analysis contains records dated from 2017 through 2025 and covers Ivano-Frankivsk together with surrounding settlements of the urban territorial community. The report was obtained from the system's accident-concentration page (https://dtp.if.ua/report).

For each accident record, the dataset contains the date and time of occurrence, settlement, street, an additional location description, accident type, and casualty outcomes. Casualties are divided into minor injuries, severe injuries, and fatalities. The report page groups individual records into spatial clusters and displays the aggregated casualty indicators and the resulting integral hazard score.

The report was generated with a spatial distance threshold of $\varepsilon = 50$m and a minimum retained cluster size of $N_{min} = 3$.

For every identified cluster C_k, the system calculated the number of accidents N_k, the aggregated numbers of minor injuries M_k, severe injuries S_k, fatalities F_k, and the integral hazard score defined in Section III as

$$H_k = M_k + 3S_k + 9F_k.$$

The clusters were then ordered in descending order of H_k. The values reported below reproduce the first ten clusters displayed by the system at the time of data extraction.

B. Cluster-level results. Table 1 presents the ten clusters with the highest integral hazard scores. Because the report does not assign descriptive identifiers to clusters, representative settlement, street, or corridor names are used to describe their locations. The final column gives the accident-frequency rank within the selected top-ten set, not within the complete set of clusters.

TABLE 1. Ten highest-ranked accident clusters according to the integral hazard score.

Hazard rank	Representative location	N_k	M_k	S_k	F_k	H_k	Frequency rank within top ten
1	Tysmenychany	36	37	8	8	133	4
2	Berezivka	20	15	11	7	111	9
3	Dragomyrchany	38	33	18	1	96	2
4	Cherniiv, Nadvirnianska St.	23	16	12	4	88	8
5	Ivano-Frankivsk, Nezalezhnosti area	27	31	8	2	73	5
6	Ivano-Frankivsk, Konovaltsia corridor	37	32	6	2	68	3
7	Ivano-Frankivsk, Halytska corridor	40	39	9	0	66	1
8	Ivano-Frankivsk, Mykolaichuka–Ivasiuka area	27	20	9	1	56	5
9	Chukalivka	13	20	4	2	50	10
10	Ivano-Frankivsk, Dovzhenka area	24	18	7	1	48	7
	Total for ten clusters	285	261	92	28	789	

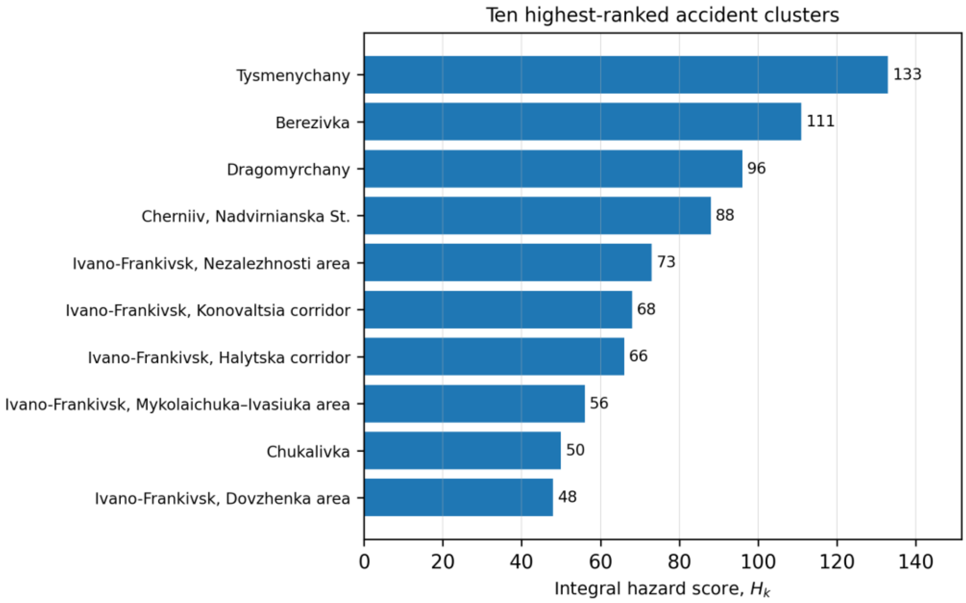

FIG. 3. Integral hazard scores of the ten highest-ranked accident clusters.

The ten highest-ranked clusters collectively contained 285 accident records. Their casualty totals were 261 minor injuries, 92 severe injuries, and 28 fatalities. Application of the selected weighting coefficients produced a cumulative score of

$$261 + 3 \times 92 + 9 \times 28 = 789.$$

The highest score was obtained for the Tysmenychany cluster. It contained 36 accident records, 37 minor injuries, eight severe injuries, and eight fatalities, producing

$$H_1 = 37 + 3 \times 8 + 9 \times 8 = 133.$$

The Berezivka cluster occupied the second position despite containing only 20 accident records. Its high score was caused by 11 severe injuries and seven fatalities:

$$H_2 = 15 + 3 \times 11 + 9 \times 7 = 111.$$

The Dragomyrchany cluster contained 38 accidents, 33 minor injuries, 18 severe injuries, and one fatality, giving $H_3 = 96$. The Cherniiv cluster contained 23 accidents but was ranked fourth because its consequences included 12 severe injuries and four fatalities, resulting in $H_4 = 88$. These results show that the highest-priority locations are not determined exclusively by the number of recorded accidents.

C. Comparison with Accident-Frequency Ranking. A comparison of the two ranking criteria demonstrates the practical effect of casualty weighting. Within the ten highest-scoring clusters, the Halytska Street corridor had the largest accident count ($N_k = 40$). However, it contained no fatalities and obtained $H_k = 66$, placing it seventh in the casualty-weighted ranking. Tysmenychany ranked fourth by accident count within the selected set but first by H_k because the cluster included eight fatalities.

Berezivka demonstrates an even larger ranking change: it ranked ninth by accident count within the top-ten set but second by H_k because of seven fatalities and 11 severe injuries. Chukalivka contained 13 accidents, whereas the Dovzhenka area contained 24; nevertheless, Chukalivka received the higher hazard score because it included two fatalities:

$$H_{Chukalivka} = 50 > H_{Dovzhenka} = 48.$$

The indicators N_k and H_k therefore characterize different aspects of road safety. N_k reflects the recorded accident frequency, whereas H_k reflects the accumulated severity of casualty outcomes. Both indicators should be retained in the analytical output rather than treated as interchangeable measures.

D. Detailed example of cluster assessment. The Nezalezhnosti Street cluster in Ivano-Frankivsk provides an illustrative example of the calculation. The cluster includes records associated with Nezalezhnosti Street and the adjacent Khryplynska and Yosyfa Slipoho streets. It contains 27 accidents with 31 minor injuries, eight severe injuries, and two fatalities. Accordingly,

$$H_k = 31 + 3 \times 8 + 9 \times 2 = 73.$$

This cluster occupies the fifth position in the weighted ranking. The underlying records include side collisions, frontal collisions, pedestrian crashes, and cyclist crashes. For example, the accident recorded on 19 December 2025 resulted in three minor injuries and two severe injuries and therefore received an individual score of nine points. Aggregation of this record with the remaining accidents produced the cluster-level value reported above.

E. Discussion. The results demonstrate that casualty-weighted ranking differs materially from ranking based solely on accident frequency. Locations with many accidents but predominantly minor outcomes may move downward, whereas clusters containing fatalities and severe injuries move upward. For intervention prioritization, this distinction prevents severe but less frequent accident concentrations from being obscured by locations with a larger number of comparatively low-severity events.

The integral score should not be interpreted as a direct estimate of accident probability, economic loss, or predictive risk. It is a relative prioritization indicator calculated from recorded casualty outcomes. Its values depend on the selected casualty weights, spatial distance threshold, minimum cluster size, observation period, completeness of the accident records, and accuracy of geocoding.

The score should also be interpreted together with its constituent indicators. Similar values of H_k may arise from different casualty structures: one cluster may contain numerous severe injuries, while another may contain fewer injuries but more fatalities. The system therefore presents Mk, Sk, Fk, and Hk separately so that analysts can examine both the integral score and the casualty structure responsible for it.

F. Infocommunication and IoT integration architecture. The developed information system can be represented as a multi-layer infocommunication architecture comprising data acquisition, communication, data management, analytical processing, and visualization layers. In the current implementation, accident records are obtained from official reporting forms and stored in a relational database after validation and geocoding.

The data acquisition layer can subsequently be extended through integration with IoT-enabled sources, including intelligent traffic counters, connected road cameras, roadside monitoring devices, weather stations, vehicle communication systems, and mobile reporting applications. These devices may provide data on traffic volume, vehicle speed, weather conditions, illumination, road surface conditions, and the presence of vulnerable road users.

The communication layer is responsible for transferring the collected information to the central processing subsystem through standardized application programming interfaces or lightweight IoT communication protocols. The received data are validated, temporally synchronized, and associated with geographic coordinates before being stored in the database.

The analytical layer applies the proposed spatial clustering and weighted casualty assessment method. In this architecture, the integral hazard score is not an isolated calculation but an analytical service that transforms heterogeneous accident-related data into ranked spatial risk indicators. The results are delivered to users through the visualization layer in the form of interactive maps, cluster reports, and hazard-level indicators.

It should be emphasized that the current implementation primarily processes official accident records. The direct acquisition of data from IoT devices represents an architectural extension and a direction for further development rather than functionality already evaluated in the present study.

CONCLUSIONS

The study developed a formal method for ranking road traffic accident clusters using spatial proximity and casualty severity. The method combines Haversine-based distance calculation, graph-based cluster formation, accident-level casualty weighting, and cluster-level score aggregation.

The weighting coefficients for minor injuries, severe injuries, and fatalities form an integral hazard indicator for individual accidents and spatial clusters. This indicator supplements the traditional accident-count metric and enables clusters with similar accident frequencies to be differentiated according to the severity of their consequences.

The method was implemented within a web-based accident analysis system that calculates cluster indicators, ranks the identified locations, and presents the results through analytical tables and map visualizations.

The results obtained can be utilized by local authorities and relevant agencies for planning road safety measures, optimizing resource allocation, and reducing accident rates. Prospects for further research include extending the hazard assessment model by incorporating additional factors, such as traffic intensity, characteristics of road infrastructure, and temporal patterns of accident occurrence.

AUTHOR CONTRIBUTIONS

B.P. – conceptualization, methodology, investigation, writing (original draft preparation), writing (review and editing).

COMPETING INTERESTS

The author declares no competing interests.

REFERENCES

[1] A. Ziakopoulos and G. Yannis, “A review of spatial approaches in road safety,” *Accident Analysis & Prevention*, vol. 135, Art. no. 105323, 2020. doi: 10.1016/j.aap.2019.105323.

[2] Z. Xie and J. Yan, “Kernel density estimation of traffic accidents in a network space,” *Computers, Environment and Urban Systems*, vol. 32, no. 5, pp. 396–406, 2008. doi: 10.1016/j.compenvurbsys.2008.05.001.

[3] T. K. Anderson, “Kernel density estimation and K-means clustering to profile road accident hotspots,” *Accident Analysis & Prevention*, vol. 41, no. 3, pp. 359–364, 2009. doi: 10.1016/j.aap.2008.12.014.

[4] S. Erdogan, I. Yilmaz, T. Baybura, and M. Gullu, “Geographical information systems aided traffic accident analysis system case study: City of Afyonkarahisar,” *Accident Analysis & Prevention*, vol. 40, no. 1, pp. 174–181, 2008. doi: 10.1016/j.aap.2007.05.004.

[5] E. S. M. M. Zahran, S. J. Tan, E. H. A. Tan, N. A. A. B. M. ’Asri Putra, Y. H. Yap, and E. K. Abdul Rahman, “Spatial analysis of road traffic accident hotspots: Evaluation and validation of recent approaches using road safety audit,” *Journal of Transportation Safety & Security*, vol. 13, no. 6, pp. 575–604, 2021. doi: 10.1080/19439962.2019.1658673.

[6] M. R. Islam, I. J. Jenny, M. Nayon, M. R. Islam, M. Amiruzzaman, and M. Abdullah-Al-Wadud, “Clustering algorithms to analyze the road traffic crashes,” in *Proc. 2021 International Conference on Science & Contemporary Technologies (ICSCT)*, 2021, pp. 1–6. doi: 10.1109/ICSCT53883.2021.9642542.

[7] M. S. Alam and N. J. Tabassum, “Spatial pattern identification and crash severity analysis of road traffic crash hot spots in Ohio,” *Heliyon*, vol. 9, no. 5, Art. no. e16303, 2023. doi: 10.1016/j.heliyon.2023.e16303.

[8] W. T. Gedamu, U. Plank-Wiedenbeck, and B. T. Wodajo, “A spatial autocorrelation analysis of road traffic crash by severity using Moran’s I spatial statistics: A comparative study of Addis Ababa and Berlin cities,” *Accident Analysis*

& *Prevention*, vol. 200, Art. no. 107535, 2024. doi: 10.1016/j.aap.2024.107535.

[9] A. Mesic, J. Damsere-Derry, C. Feldacker, S. J. Mooney, A. Gyedu, C. Mock, A. Kitali, B. H. Wagenaar, et al., “Identifying emerging hot spots of road traffic injury severity using spatiotemporal methods: Longitudinal analyses on major roads in Ghana from 2005 to 2020,” *BMC Public Health*, vol. 24, Art. no. 1609, 2024. doi: 10.1186/s12889-024-18915-x.

[10] S. Mohammed, A. H. Alkhereibi, A. Abulibdeh, R. N. Jawarneh, and P. Balakrishnan, “GIS-based spatiotemporal analysis for road traffic crashes: In support of sustainable transportation planning,” *Transportation Research Interdisciplinary Perspectives*, vol. 20, Art. no. 100836, 2023. doi: 10.1016/j.trip.2023.100836.

[11] Y. Berhanu, D. Schröder, B. T. Wodajo, and E. Alemayehu, “Machine learning for predictions of road traffic accidents and spatial network analysis for safe routing on accident and congestion-prone road networks,” *Results in Engineering*, vol. 23, Art. no. 102737, 2024. doi: 10.1016/j.rineng.2024.102737.

[12] M. Tahir, Y. Qiao, N. Kanwal, B. Lee, and M. N. Asghar, “Real-time event-driven road traffic monitoring system using CCTV video analytics,” *IEEE Access*, vol. 11, pp. 139097–139111, 2023. doi: 10.1109/ACCESS.2023.3340144.

[13] D. K. Endashaw, K. A. Habtegiorgis, B. M. Al-Ramadan, H. M. Al-Ahmadi, and B. F. Deressa, “A systematic review on GIS-based road traffic accidents analysis and road safety audit,” *Computational Urban Science*, vol. 5, Art. no. 53, 2025. doi: 10.1007/s43762-025-00221-w.

[14] B. Pashkovskyi and M. Slabinoha, “Automation of determination of areas and places of concentration of traffic accidents at the Ivano-Frankivsk urban territorial community using geocoding,” *International Science Journal of Engineering & Agriculture*, vol. 2, no. 4, pp. 9–18, 2023. doi: 10.46299/j.isjea.20230204.02.

[15] B. V. Pashkovskyi and M. I. Slabinoha, “Improvement of the automated process for identifying road traffic accident concentration areas and locations in the Ivano-Frankivsk urban territorial community,” *Bulletin of Kherson National Technical University*, no. 2(89), pp. 204–210, 2024. doi: 10.35546/kntu2078-4481.2024.2.29.

[16] B. V. Pashkovskyi, “Automated process for identifying road traffic accident concentration areas and locations in the Ivano-Frankivsk urban territorial community,” *Bulletin of Kherson National Technical University*, no. 2(93), pp. 288–292, 2025. doi: 10.35546/kntu2078-4481.2025.2.2.34.

Bohdan Pashkovskyi

PhD (Tech.), Associate Professor at the Department of Computer Systems and Networks, Ivano-Frankivsk National Technical University of Oil and Gas. His research focuses on optimal control of gas compression under uncertainty. Author of 15+ publications, including Scopus-indexed papers. Supervisor of student research and software engineer with industry experience.

ORCID ID: 0000-0003-1082-6837

Інтегральна оцінка небезпечності ділянок концентрації дорожньо-транспортних пригод на основі вагових коефіцієнтів постраждалих

Богдан Пашковський*

Кафедра комп’ютерних систем і мереж, Івано-Франківський національний технічний університет нафти і газу, Івано-Франківськ, Україна

*Автор-кореспондент (Електронна адреса: bohdan.pashkovskyi@nung.edu.ua)

АНОТАЦІЯ У статті розглянуто задачу формалізованого ранжування ділянок концентрації дорожньо-транспортних пригод з урахуванням не лише кількості зафіксованих подій, а й тяжкості їх наслідків. Актуальність дослідження зумовлена тим, що традиційне впорядкування аварійно-небезпечних ділянок за частотою ДТП може недооцінювати локації з меншою кількістю пригод, але значною кількістю важко травмованих і загиблих. Запропоновано метод інтегральної оцінки небезпечності, у якому для кожної ДТП обчислюється зважений показник на основі кількості легко травмованих, важко травмованих і загиблих. Для цих категорій застосовано вагові коефіцієнти 1, 3 і 9 відповідно. На рівні просторового кластера індивідуальні оцінки агрегуються в інтегральний показник, що дає змогу одночасно враховувати частоту ДТП та тяжкість їх наслідків. Просторове групування подій виконується за географічними координатами. Відстань між точками визначається за формулою гаверсинуса, а множина ДТП подається у вигляді неорієнтованого графа близькості. Ребро між двома вершинами формується, якщо відстань між відповідними ДТП не перевищує заданого порогового значення. Кластери визначаються як компоненти зв’язності цього графа. Запропонований алгоритм охоплює валідацію та геокодування записів, обчислення індивідуальних оцінок, попарне порівняння координат, формування кластерів, агрегування показників постраждалих і ранжування кластерів за інтегральною оцінкою. Обчислювальна складність базової реалізації становить $O(n)^2$), оскільки домінуючою операцією є попарне порівняння записів. Метод реалізовано у складі веборієнтованої інформаційної системи аналізу ДТП. Експериментальне оцінювання виконано на записах за 2017–2025 роки для Івано-Франківської міської територіальної громади з порогом просторової близькості 50 м і мінімальним розміром кластера 3. Для кожного кластера система визначає кількість ДТП, сумарну кількість легко травмованих, важко травмованих і загиблих, а також інтегральний показник небезпечності. Аналіз десяти кластерів із найвищими оцінками показав, що ранжування за тяжкістю наслідків істотно відрізняється від ранжування лише за кількістю ДТП. Зокрема, кластери з меншою кількістю подій, але більшою кількістю загиблих і важко травмованих, отримують вищий пріоритет. Запропонований показник не є прямою оцінкою ймовірності ДТП або економічних збитків, а використовується як відносний інструмент пріоритизації. Розроблений метод може бути аналітичним компонентом інфокомунікаційної системи моніторингу безпеки дорожнього руху та надалі розширюватися шляхом інтеграції з IoT-пристроями дорожньої інфраструктури.

КЛЮЧОВІ СЛОВА дорожньо-транспортна пригода, кластер ДТП, інтегральна оцінка небезпечності, кластеризація, безпека дорожнього руху.