

Appendix A. Processes of HDT trip chain enhancement

This step enhances the accuracy and reliability of HDT trip chain identification (**Fig. A.1**). First, for the trip chain identified in Step 3 of Section 3.1 (**Fig. A.1a**), trajectory densification and sparsification techniques are applied to mitigate trajectory discontinuities and GPS drifts by adjusting the spatial and temporal resolution of trajectory points (**Fig. A.1b**). These issues often arise when HDTs operate in dense urban areas, passing under bridges, or idling near large infrastructure. Next, a trajectory smoothing process using Kalman filtering, as proposed by Liu and Yang (2023), is applied to reduce fluctuations in GPS traces (**Fig. A.1c**). This enhances the consistency of vehicle movement patterns, providing a cleaner and more reliable representation of HDT trajectories. Subsequently, a map-matching algorithm is employed to align raw GPS points with the underlying road network, ensuring that the trip chains accurately correspond to real-world routes and freight activity (**Fig. A.1d**). The enhanced trip chains (**Fig. A.1e**) are then represented as time series, serving as inputs for the MRFM which is introduced in the next subsection.

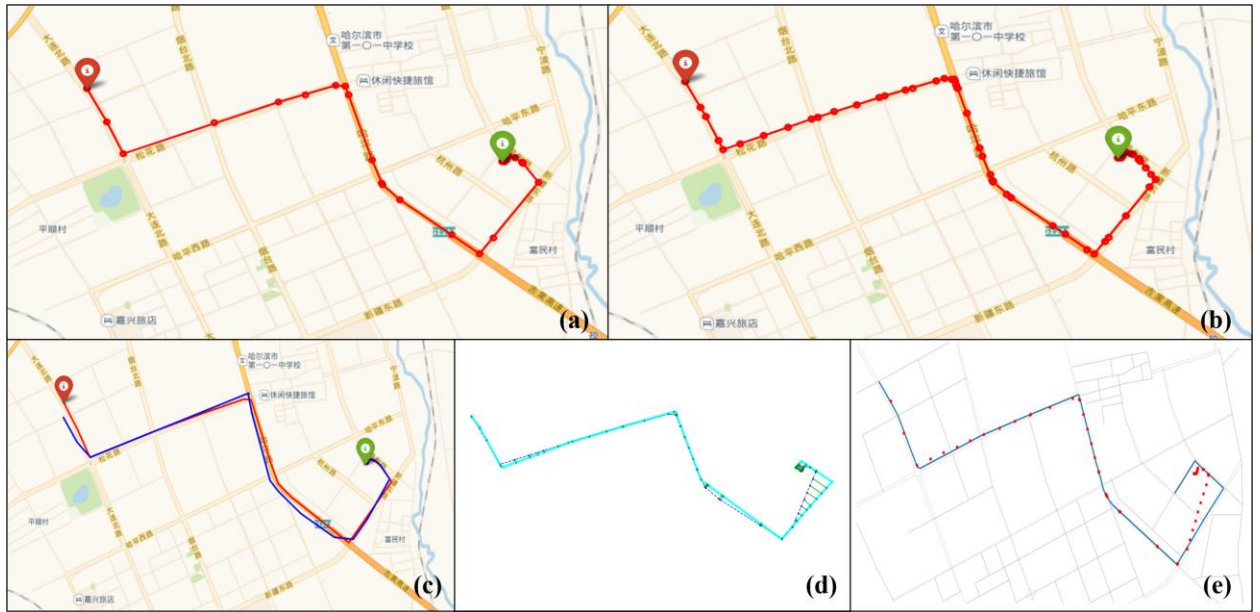


Fig. A.1. HDT Trip Chain Enhancement: (a) Trip chain identified in Step 3; (b) Densified or sparsified trip chain; (c) Smoothed trip chain; (d) Trip chain after map matching; (e) Enhanced trip chain for MRFM.

In trajectory smoothing, each data point represents an observation of the vehicle's actual state.

However, measurement errors often introduce deviations from the true position. To correct these discrepancies, the process begins with drift removal, followed by Kalman filtering to refine state estimation. The Kalman filter combines predicted positions with current observations, reducing measurement noise and mitigating minor fluctuations—particularly under conditions of stable noise variance. It operates in two primary stages: prediction and correction. In the prediction step, as defined in **Eq. (A.1)**, prior data is used to forecast the vehicle’s state at time k , denoted as $\hat{X}_{k|k-1}$. The subsequent correction step improves this prediction by incorporating new observations, resulting in an updated state estimate, $\hat{X}_{k|k}$. The core Kalman filter equations, provided in **Eq. (A.2)** and **Eq. (A.3)**, describe the Time Update (prediction) and Measurement Update (correction) processes. These steps involve the state transition matrix F_k , measurement matrix H_k , process noise covariance Q_k , measurement noise covariance R_k , control input matrix B_k , and z_k , representing the observation equation. This iterative process integrates prior knowledge with real-time data to yield an accurate and reliable estimate of the HDT’s trajectory.

$$\hat{X}_{k-1|k-1} \xrightarrow{\text{Prediction}} \hat{X}_{k|k-1} \xrightarrow{\text{Correction}} \hat{X}_{k|k} \quad (\text{A.1})$$

$$\begin{cases} \hat{x}_{k|k-1} = F_k \hat{x}_{k-1} + B_k u_k \\ P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k \end{cases} \quad (\text{A.2})$$

$$\begin{cases} K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1} \\ \hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (z_k - H_k \hat{x}_{k|k-1}) \\ P_{k|k} = (I - K_k H_k) P_{k|k-1} \end{cases} \quad (\text{A.3})$$

Following trajectory smoothing, Distance-Based Map Matching is applied to align trajectory points with the road network by selecting the closest road segments based on a predefined distance threshold and calculating the distances to the nearest road segments. The map matching procedure consists of two distinct phases: a prediction phase, where the trajectory is initially mapped to the road network, and a correction phase, during which the match is refined using real-time observations. To enhance map matching accuracy, factors such as observation noise, the initial matching distance, and temporal consistency are considered, thereby improving the reliability and precision of the road network mapping.

Appendix B. Details for multi-feature extraction

To extensively evaluate the characteristics of HDT trip chains at various resolutions, this study

considers both trip-related (including spatial, temporal, and link features) and activity-related characteristics (freight-related features). For each freight-related stop p_k within the trip chain t_q , a total of 14 distinct features are considered. **Eqs. (B.1)-(B.17)** represent the corresponding features under R_m , which are defined by the functions f_1^m through f_{14}^m , with the variable definitions provided in **Table B.1**.

Table B.1

Symbols and definitions.

Symbol	Description
$g_{i,j}^m$	A hexagonal grid cell at the m -th resolution level, indexed by (i, j) .
r_m	The side length of a hexagonal grid cell at the m -th resolution level.
ω_m	The weight associated with the m -th resolution level R_m .
F^m	The feature matrix at the m -th resolution level R_m .
F	The feature matrix across all resolution levels $\{R_m\}_{m=1}^M$.
(x, y)	The geographical latitude and longitude coordinates of point p .
(i, j)	The grid index corresponding to the coordinates of point p .
G_i	Coordinates of the i -th GPS data point.
g_j	Coordinates of the j -th grid center.
$R(\theta)$	The rotation angle θ used to align the hexagonal grid with the GPS trajectory.
$N_{i,j}^m$	The number of logistics nodes in grid $g_{i,j}^m$ at the m -th level resolution.
$P_{i,j}^m(\tau)$	The proportion of logistics nodes in grid $g_{i,j}^m$ during each period $\tau = \{\tau_1, \dots, \tau_{24}\}$.
Δt_n	The stay duration of the n -th logistics node.
$1(\cdot)$	Indicator function.
U	The upstream point set.
D	The downstream point set.
N_U^m	The number of upstream GPS trajectory points at the m -th level resolution.
N_D^m	The number of downstream GPS trajectory points at the m -th level resolution.
θ_r	The upstream direction angle of the r -th upstream point.
θ_q	The downstream direction angle of the q -th downstream point.
$p_{i,j}^m(\theta)$	The primary direction distribution density, represented by Kernel Density Estimation (KDE).
$d_H(\cdot)$	Haversine distance function.
$ P_k $	The distance from point P_k to the center of the sphere.
ω_k	The weight of POI K , based on the proportion of the category to which the POI belongs.
d_k	The distance from POI K to its nearest logistics node.
$\varphi(\cdot)$	The distance decay function, where β represents a rate parameter controlling the decay rate.

(1) Trip-related characteristics

Spatial features, including geographical coordinates (**Eq. (B.1)**), grid indices at varying resolutions (**Eq. (B.2)**), and logistics node density (**Eq. (B.3)**), are essential for effectively segmenting the study area and ensuring that the chosen resolution adequately captures the spatial heterogeneity inherent in the data.

$$f_1^m(p) = x, f_2^m(p) = y \quad (\text{B.1})$$

$$f_3^m(p) = i, f_4^m(p) = j \quad (\text{B.2})$$

$$f_5^m(p) = \frac{N_{i,j}^m}{A_{i,j}^m} \quad (\text{B.3})$$

Temporal features, such as the stay duration at each point (**Eq. (B.4)**), peak period within grids (**Eqs. (B.5)–(B.6)**), and average stay time within grids (**Eq. (B.7)**), are crucial for understanding the spatiotemporal dynamics of traffic patterns and identifying temporal congestion.

$$f_6^m(p) = \Delta t \quad (\text{B.4})$$

$$P_{i,j}^m(\tau) = \frac{\sum_{n=1}^{N_{i,j}^m} 1(\tau_n \in \tau)}{N_{i,j}^m} \quad (\text{B.5})$$

$$f_7^m(p) = \arg \max_t P_{i,j}^m(\tau) \quad (\text{B.6})$$

$$f_8^m(p) = \overline{\Delta t_{i,j}^m} = \frac{\sum_{n=1}^{N_{i,j}^m} \Delta t_n}{N_{i,j}^m} \quad (\text{B.7})$$

Key link features, such as traffic flow, grid density, upstream and downstream directionality, and average speed, are crucial for capturing dynamic travel patterns in trip chains. These factors comprehensively describe traffic movement and congestion across network links. Let the upstream point set $U = \{u_r \mid t(u_r) < t(p_{first}), r = 1, 2, \dots, N_U\}$ represent all points prior to the first logistics node $p_{first} = (x_{first}, y_{first})$ within the grid, and the downstream point set $D = \{u_q \mid t(u_q) > t(p_{last}), q = 1, 2, \dots, N_D\}$ denote all points following the last logistics node $p_{last} = (x_{last}, y_{last})$. The metrics for traffic flow and grid density, as defined in **Eq. (B.8)** and **Eq. (B.9)**, quantify the number of GPS points passing through grid cell $g_{i,j}^m$ per unit time, and the number of upstream and downstream GPS points per unit area within the defined grid boundaries. These characteristics are fundamental for optimizing HDT operations by providing insights into congestion patterns, route efficiency, and resource allocation.

$$f_9^m(p) = \frac{N_U^m + N_D^m}{t(p_{last}) - t(p_{first})} \quad (\text{B.8})$$

$$f_{10}^m(p) = \frac{N_U^m + N_D^m}{A_{i,j}^m} \quad (\text{B.9})$$

Eqs. (B.10)–(B.14) define the dominant directions, as proposed by Hsueh and Chen (2018). **Eq. (B.10)** and **Eq. (B.11)** calculate the direction angles for upstream and downstream links, respectively, while **Eq. (B.12)** employs KDE to quantify the distribution of dominant directions. **Eq. (B.13)** and **Eq. (B.14)** identify the dominant directions within these upstream and downstream links, corresponding to the maximum values of the direction angle density. Understanding these dominant directions is essential for accurately predicting the trajectory of HDT trip chains, optimizing route selection, reducing congestion-induced delays, and ultimately designing more efficient travel corridors.

$$\theta_r = \arctan \left(\frac{\sin(x_{first} - \text{long}_{u_r}) \cdot \cos(y_{first})}{\cos(\text{lat}_{u_r}) \cdot \sin(y_{first}) - \sin(\text{lat}_{u_r}) \cdot \cos(\text{lat}_{u_r}) \cdot \cos(x_{first} - \text{long}_{u_r})} \right) \quad (\text{B.10})$$

$$\theta_q = \arctan \left(\frac{\sin(\text{long}_{d_q} - x_{last}) \cdot \cos(\text{lat}_{d_q})}{\cos(y_{last}) \cdot \sin(\text{lat}_{d_q}) - \sin(y_{last}) \cdot \cos(\text{lat}_{d_q}) \cdot \cos(\text{long}_{d_q} - x)} \right) \quad (\text{B.11})$$

$$P_{i,j}^m(\theta) = \frac{1}{(N_U^m + N_D^m)h} \sum_{k=1}^N \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{(\theta - \theta_{k \in \{U \cup D\}})^2}{2h^2} \right) \quad (\text{B.12})$$

$$f_{11}^m(p) = \arg \max_{\theta} P_{i,j}^m(\theta_{r \in U}) \quad (\text{B.13})$$

$$f_{12}^m(p) = \arg \max_{\theta} P_{i,j}^m(\theta_{q \in D}) \quad (\text{B.14})$$

The average speed in $g_{i,j}^m$ at the m -th level of resolution is defined as the mean speed from the upstream point to the first stop and from the last stop to the downstream point (**Eq. (B.15)**). The speed along the link serves as a key indicator of traffic flow dynamics, providing valuable insights for assessing the impact of traffic conditions and congestion, ultimately informing decisions related to HDT schedule planning.

$$f_{13}^m(p) = \frac{1}{N_U^m + N_D^m} \sum_{r=1}^{N_U} \left(\frac{d_H(\text{long}_{u_r}, \text{lat}_{u_r}, x_{first}, y_{first})}{|t_{u_r} - t_{first}|} \right) + \sum_{q=1}^{N_D} \left(\frac{d_H(\text{long}_{d_q}, \text{lat}_{d_q}, x_{last}, y_{last})}{|t_{last} - t_{d_q}|} \right) \quad (\text{B.15})$$

$$d_H = 2R \cdot \arcsin \left(\sqrt{\sin^2 \left(\frac{\Delta P}{2} \right) + \|P_1\| \cdot \|P_2\| \cdot \sin^2 \left(\frac{\Delta \lambda}{2} \right)} \right) \quad (\text{B.16})$$

(2) Activity-related characteristics

Freight-related features, particularly the density of freight-related POIs within each grid cell (Eq. (B.17)), are essential for identifying high-traffic zones and optimizing the placement of logistics infrastructure. By incorporating the semantic characteristics of POIs and their proximity to stops, the accuracy of trip chain identification for HDTs is improved, thereby enhancing the ability to develop more tailored HDT transportation plans.

$$f_{14}^m(p) = \frac{\sum_{k \in POI_K} \omega_k \cdot \varphi(d_k)}{A_{l,j}^m} \quad (\text{B.17})$$

$$\varphi(d_k) = e^{-\beta d_k} \quad (\text{B.18})$$

Appendix C. Model validation and evaluation results

This section evaluates the performance of the proposed MRSF against five single-resolution baselines. The fused resolution is denoted as r , while five single-resolution levels (250 m, 353 m, 500 m, 707 m, and 1000 m), denoted as r_1 to r_5 , were selected using the method described in Section 3.2.1. For each level, a distance matrix was constructed to quantify trip chain similarity under a uniform spatial scale. These single-resolution results were then compared against the fused multi-resolution output. The evaluation spans clustering performance, classification accuracy, dimensionality reduction effectiveness, visualization quality, and pattern stability and robustness. Optimal values are highlighted in bold in **Table C.1**.

Clustering: Five commonly used clustering algorithms, namely K-means, DBSCAN, Spectral Clustering, GMM, and Mini-Batch K-means, were applied to the similarity matrices derived from both MRSF and the single-resolution baselines. The silhouette score was used as the primary evaluation metric, as it effectively measures the compactness and separability of clusters based on distance metrics. As shown in **Table C.1**, the results demonstrate that MRSF consistently outperforms all single-resolution models across all clustering algorithms.

Table C.1

Silhouette scores of clustering algorithms at different resolutions.

Method	r	r_1	r_2	r_3	r_4	r_5
K-means	0.1322	0.1141	0.1148	0.1140	0.1191	0.1214
DBSCAN	0.3799	0.3171	0.3033	0.3223	0.3096	0.3236
Spectral clustering	0.2155	0.1660	0.1348	0.1963	0.1787	0.1971

GMM	0.2472	0.2147	0.2101	0.1179	0.1215	0.1816
Mini-batch K-means	0.3383	0.1052	0.1771	0.2213	0.1098	0.2031

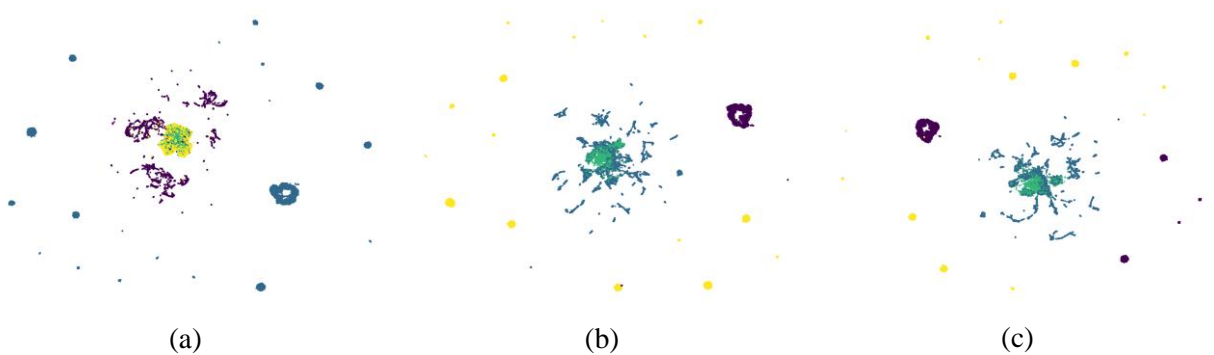
Classification: a One-Class Support Vector Machine (SVM) classifier was applied to the similarity matrices generated from both the multi-resolution fusion and the single-resolution baselines. The silhouette score was again used as the evaluation metric to quantify the quality of classification boundaries formed based on trip chain similarities. The results in Table C.2 indicate that MRSF uniformly outperforms all individual single-resolution models.

Table C.2

Silhouette scores of One-Class SVM classification across different resolutions.

Method	r	r_1	r_2	r_3	r_4	r_5
One-Class SVM	0.2921	0.2241	0.2765	0.2798	0.1920	0.2064

Dimensionality Reduction and Visualization: Uniform Manifold Approximation and Projection (UMAP) was employed for dimensionality reduction and visualization to qualitatively assess the spatial separability of the trip chain representations. **Fig. C.–C.5** presents the clustering results based on the Mini-Batch K-means algorithm, K-means, DBSCAN, spectral clustering, and GMM, where panel (a) illustrates the fused resolution r , and panels (b) to (e) correspond to the single-resolution levels r_1 to r_5 , respectively. The visualization clearly demonstrates that MRSF substantially enhances cluster compactness and separability at the fusion level compared to single-resolution representations, indicating stronger spatial discrimination in the extracted features.



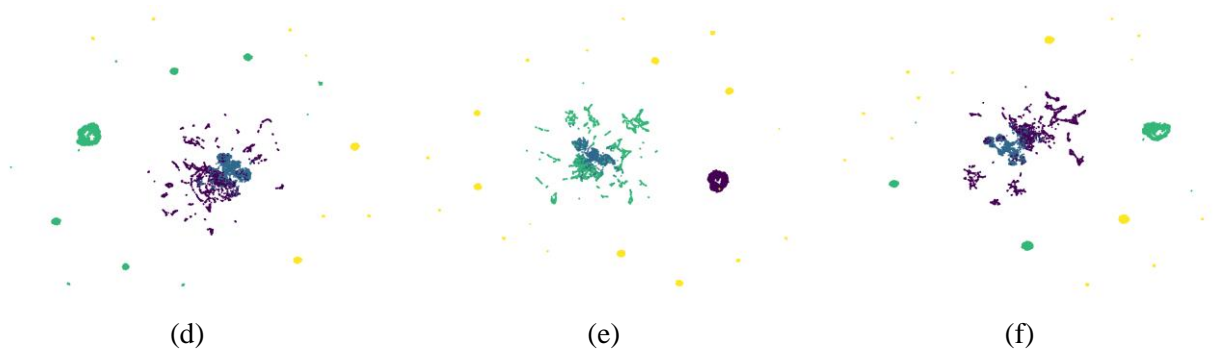


Fig. C. 1. Mini-Batch K-means clustering results across different resolutions.

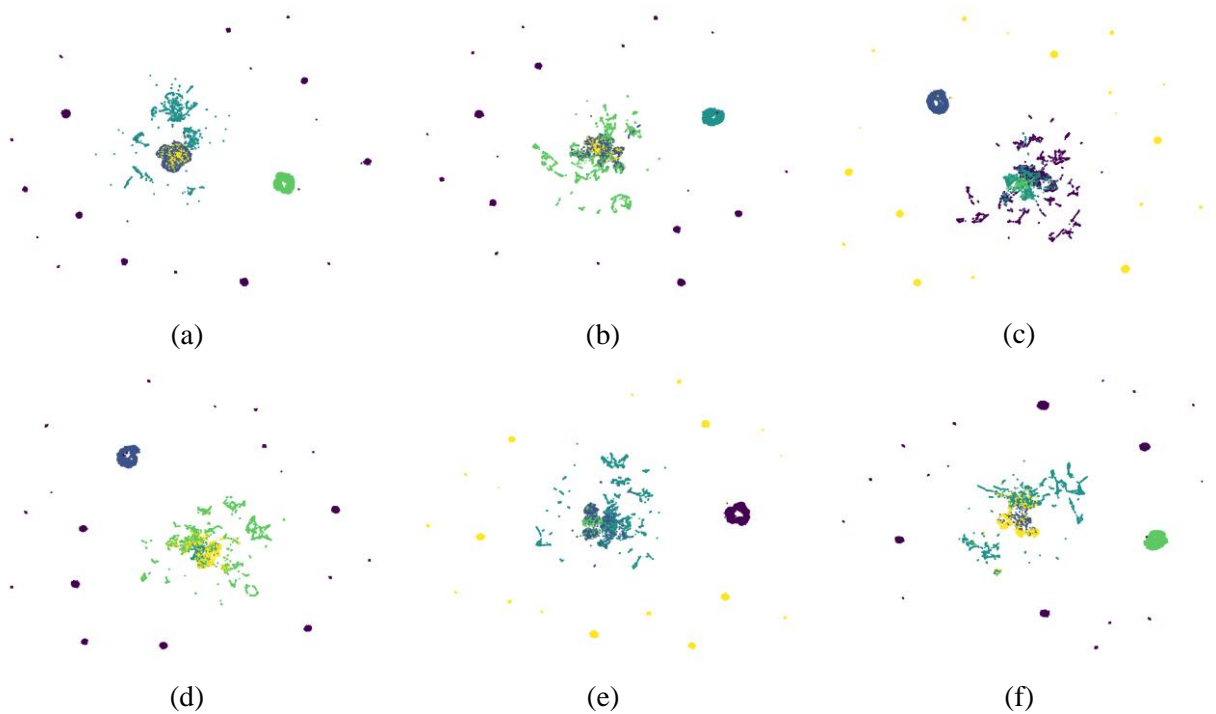
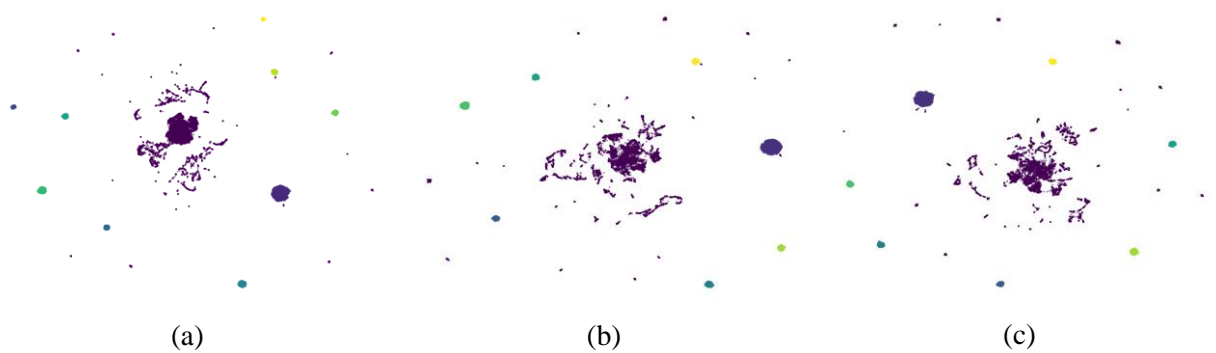


Fig. C.2. K-means clustering results across different resolutions.



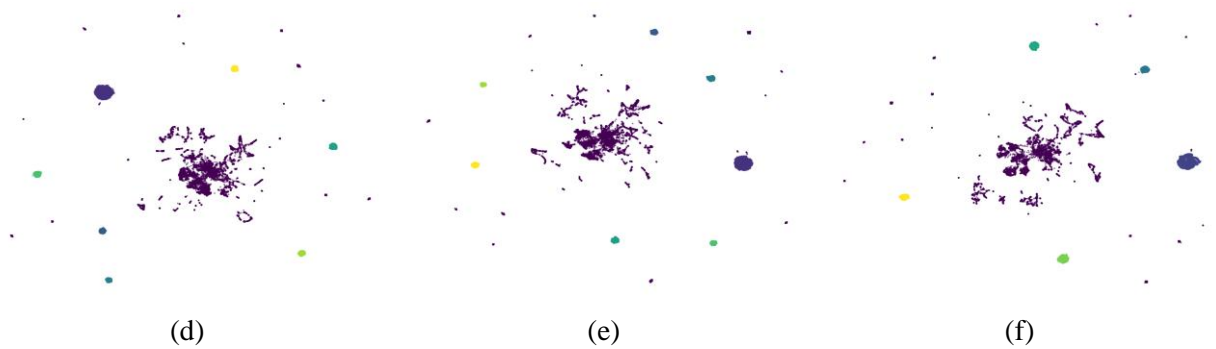


Fig. C.3. DBSCAN clustering results across different resolutions.

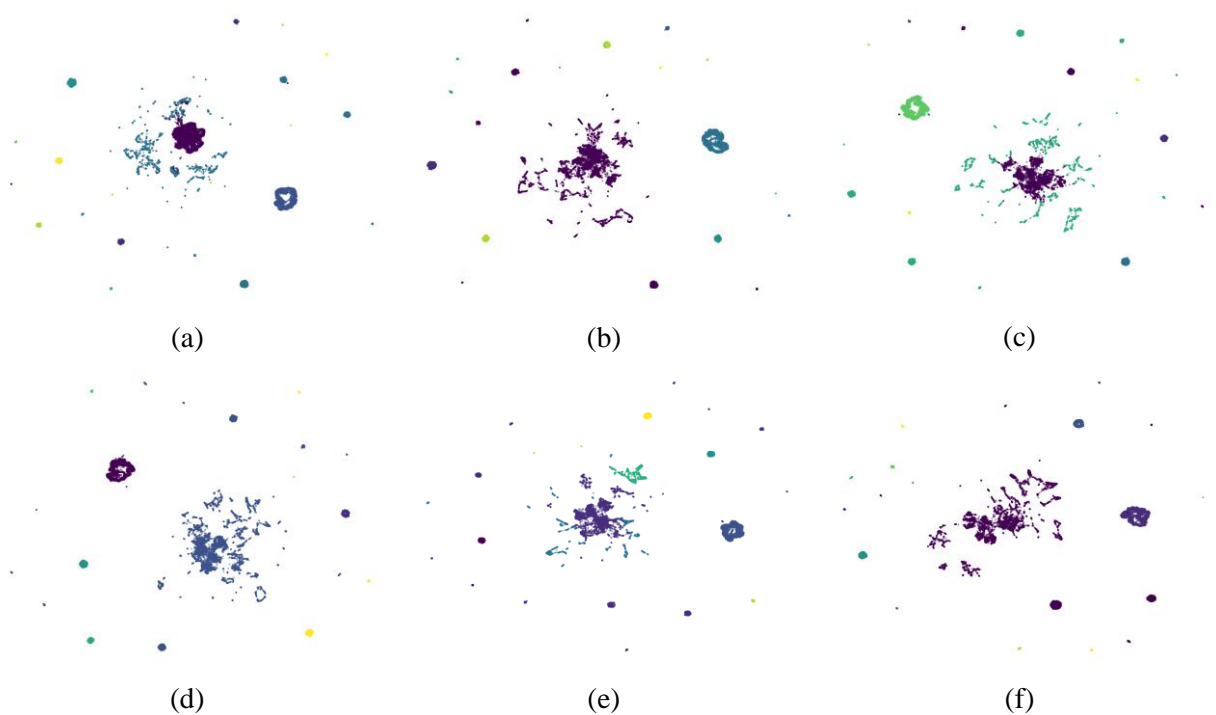
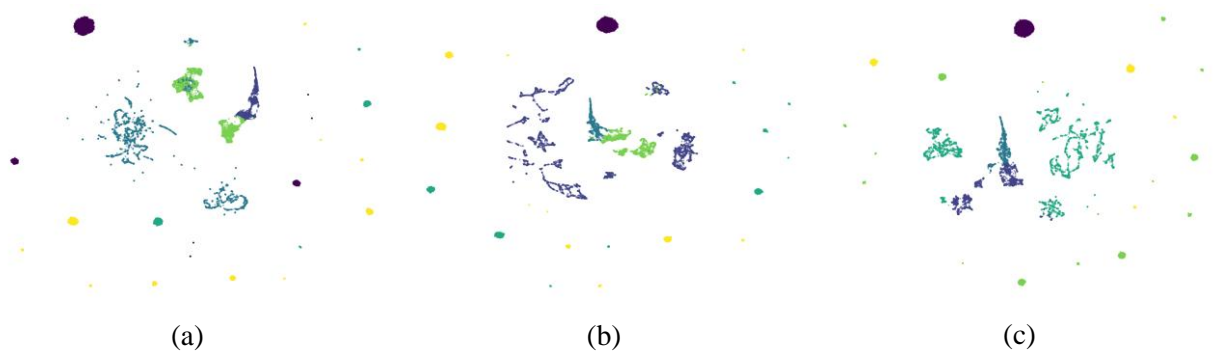


Fig. C.4. Spectral clustering results across different resolutions.



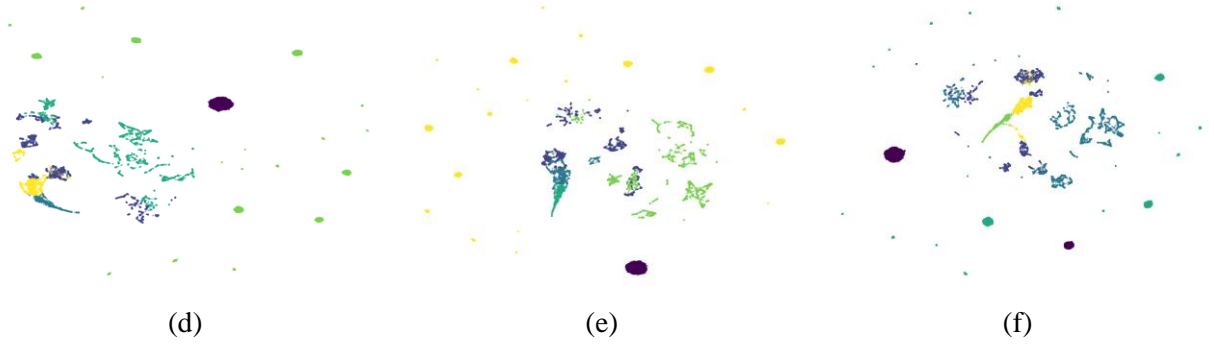


Fig. C.5. GMM clustering results across different resolutions.

Stability and Robustness: a bootstrap resampling method was employed. For each resolution level and the fusion layer, repeated sampling and re-evaluation were conducted to examine the consistency of trip chain similarity measurements. **Table C.3** summarizes the results, indicating that the fusion layer achieves the smallest mean MD-DTW distance and the lowest standard deviation among all scenarios, thereby demonstrating superior stability and robustness.

Table C.3

Bootstrap evaluation results of classification performance across different resolutions.

Method	r	r_1	r_2	r_3	r_4	r_5
Mean	567.5142	569.4981	653.5729	581.9743	577.3047	574.4655
Std	1.3937	1.4844	1.6519	1.6058	1.4254	1.4274

Overall, the validation results consistently demonstrate the advantages of the proposed MRSF over single-resolution baselines across multiple evaluation dimensions. The fused multi-resolution features significantly improve clustering quality, classification performance, and visualization interpretability, while also enhancing stability and robustness under bootstrap resampling. By integrating multi-scale spatial granularity, MRSF offers a more accurate, reliable, and comprehensive representation of HDT activity-trip chain patterns. These findings highlight both the effectiveness and generalizability of MRSF, providing a strong foundation for its application to freight pattern analysis and data-driven logistics management.

Appendix D. Process of HDT stop identification

Each GPS point is classified as stationary or moving based on its average speed. When the average

speeds between consecutive GPS points are computed and visualized on a double logarithmic scale, the resulting distribution exhibits a distinct bimodal pattern. The right peak corresponds to typical HDT travel speeds, while the left peak is primarily due to GPS drift, which creates the illusion of movement in stationary vehicles. To accurately differentiate between these two states, a two-component mixture model is employed (Eq. (D.1)), where $f(\cdot)$ denotes the overall probability density function. This model integrates a lognormal distribution $Lognorm(\cdot)$ (Eq. (D.2)) to characterize GPS-induced drift speeds and a normal distribution $Norm(\cdot)$ (Eq. (D.3)) to represent actual HDT movement. The weights and parameters of the mixture— ω_1, ω_2 for the distribution contributions, and $\mu_1, \mu_2, \sigma_1, \sigma_2$ for their respective means and standard deviations—are estimated using maximum likelihood methods.

$$f(x; \omega_1, \mu_1, \sigma_1, \omega_2, \mu_2, \sigma_2) = \omega_1 \cdot Lognorm(x; \omega_1, \mu_1, \sigma_1) + \omega_2 \cdot Norm(x; \omega_2, \mu_2, \sigma_2) \quad (D.1)$$

$$Lognorm(x; \omega_1, \mu_1, \sigma_1) = \frac{1}{x\sigma_1\sqrt{2\pi}} \exp\left(-\frac{(\ln(x) - \mu_1)^2}{2\sigma_1^2}\right) \quad (D.2)$$

$$Norm(x; \omega_2, \mu_2, \sigma_2) = \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left(-\frac{(x - \mu_2)^2}{2\sigma_2^2}\right) \quad (D.3)$$

As a result, in Harbin, an HDT is classified as stationary if its speed remains below 2.85 km/h—the saddle point identified in Fig. D.1a—for a predefined duration. The corresponding stop location is represented by the centroid of GPS points recorded during consecutive stationary intervals (Fig. D.1b).

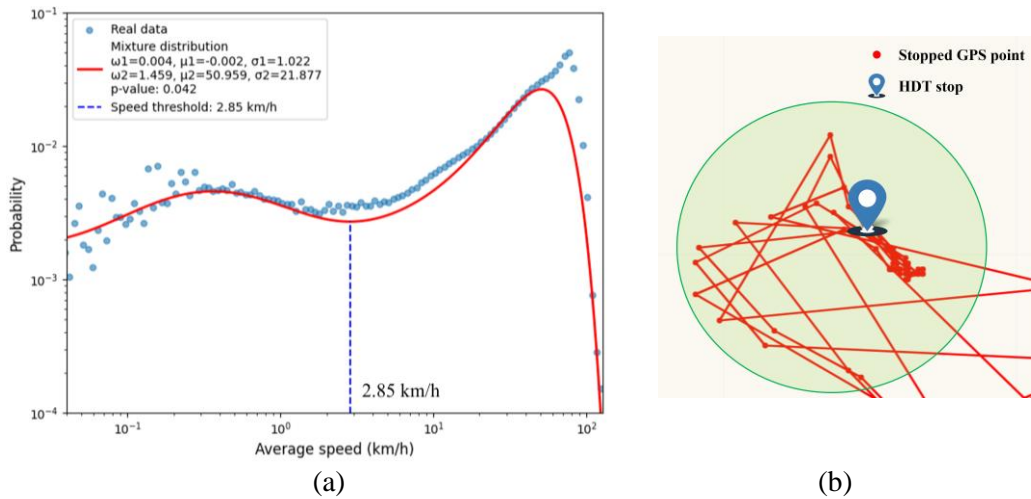


Fig. D.1. (a) Distributions of average speed in Harbin. (b) HDT stop location.

Next, the Loubar method is employed to derive multilevel thresholds for stop durations. Stop points are first ranked in descending order based on their durations. Cumulative durations are then computed and normalized to construct a Lorenz curve. The critical threshold F^* is defined as the point where the tangent to the Lorenz curve at (1,1) intersects the x-axis. Stops with normalized ranks exceeding F^* form the first group, and this value serves as the initial threshold. These points are subsequently excluded, and the process is repeated iteratively, recomputing the Lorenz curve for the remaining stops to determine successive thresholds. The iteration continues until the Lorenz curve approaches linearity, indicating a more uniform distribution. Finally, as illustrated in **Fig. D.2**, after 12 iterations, the stop durations are classified into five types: Type I (>11.2 hours), Type II (5.8–11.2 hours), Type III (2.5–5.8 hours), Type IV (37.2 minutes–2.5 hours), and Type V (7.5–37.2 minutes), reflecting a refined stratification of freight-related stop behaviors.

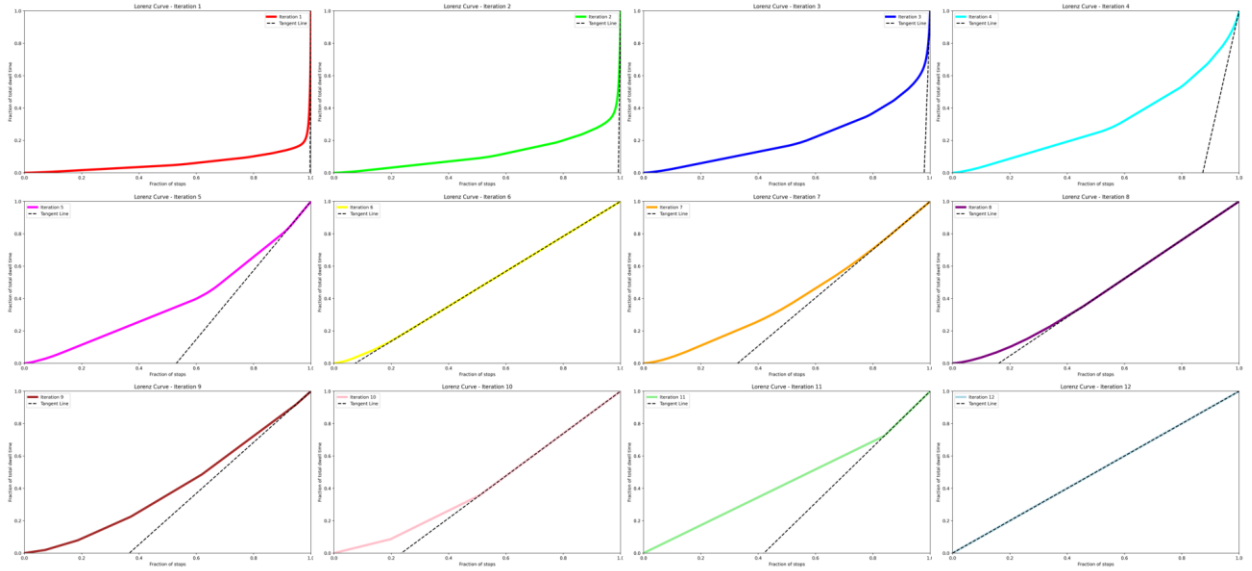


Fig. D.2. Multilevel time thresholds classification for HDT stops in Harbin (12 Iterative Phases).

Appendix E. Spectral clustering process of identifying HDT activity-trip chain patterns

In this study, five single-resolution feature matrices were derived and fused into a unified 160983×14 matrix representing all identified HDT trip chains. To accommodate the non-linear relationships and

complex cluster boundaries inherent in HDT GPS data, spectral clustering was applied to the MD-DTW-based distance matrix at the fusion level.

Each trip chain was treated as a node in a similarity graph constructed using the MD-DTW matrix, where edge weights encoded pairwise similarities. A Gaussian kernel function was applied to transform distances into similarity scores, as defined in **Eq. (E.1)**:

$$w_{ij} = \exp\left(-\frac{MDDTW(i,j)^2}{2\sigma^2}\right) \quad (\text{E.1})$$

The normalized Laplacian matrix L_{sym} was then computed by:

$$L = D - W$$

$$L_{sym} = D^{-1/2} L D^{-1/2} = \frac{L_{ij}}{\sqrt{d_i d_j}} \quad (\text{E.2})$$

where D was the degree matrix with $d_i = \sum_j w_{ij}$ and W was the weighted adjacency matrix. Spectral decomposition was performed on L_{sym} , and the smallest k_1 eigenvectors u_1, \dots, u_{k_1} were selected to form the matrix $U = [u_{ij}]_{n \times k_1} \in \mathbb{R}^{n \times k_1}$. The rows of U were normalized to unit length, yielding matrix T . Each normalized row $y_i \in \mathbb{R}^{k_1}$ represented a sample in the embedded space and was subsequently clustered using the k-means algorithm to produce clusters C_1, C_2, \dots, C_{k_2} . The final cluster assignment was defined as $A_i = \{j | \{y_j \in C_i\}, i = 1, \dots, k_2\}$.

To determine the optimal embedding dimensionality k_1 , we used the explained variance of eigenvalues. The eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of L_{sym} were sorted in ascending order, and k_1 was chosen such that the cumulative energy exceeds 90%, as calculated in **Eq. (E.3)**:

$$Cumulative\ Energy(k) = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^n \lambda_i} \quad (\text{E.3})$$

here, $k_1 = 10$. The number of clusters k_2 was determined by maximizing the silhouette coefficient, yielding an optimal value of $k_2 = 5$.

Compared to traditional clustering methods, spectral clustering is particularly effective for HDT trip chains with high-dimensional, non-convex feature distributions. Unlike DBSCAN, which is sensitive to parameters and lacks robustness for trajectory-shaped data, or k-means, which assumes convexity and suffers from local optima, spectral clustering captures intricate feature interactions and delivers consistent,

reliable cluster outcomes. Applied to the 160983×160983 MD-DTW matrix at the fusion level, the algorithm reveals five distinct activity-trip chain patterns with strong clustering validity.