
Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org

Volume 25, Issues 5, pp. 4095–4112, DOI 10.12694/scpe.v25i5.3185

RESEARCH ON DATA-DRIVEN URBAN INTELLIGENT MONITORING AND OLD CITY RECONSTRUCTION

YI WANG*

Abstract. As the global urbanization process is further accelerating, the number of urban people is also steadily increasing. With the continuous growth of the urban population, the use of various functional facilities in the city is gradually becoming saturated, which seriously restricts the development of the city. Therefore, real-time perception and prediction of the operational status of various functional urban facilities, as well as environmental safety monitoring in the city, are of great significance for improving the functionality and livability of the city. In complex urban environments, multi-source data is interrelated, and the noise reduction of multi-source data and the integration of this correlation still face great challenges. At the same time, how to apply urban intelligent monitoring in the reconstruction of old cities is rarely mentioned. Based on the above problems, this paper proposes a data noise reduction model based on a wavelet algorithm and combines it with the Macroscopic Fundamental Diagram (MFD) multi-source data fusion structure proposed in this paper to provide a theoretical basis for the construction of an urban intelligent monitoring model. Then, the author constructed a data-driven urban safety environment monitoring model and a multi-source data-based urban congestion monitoring model, which have good experimental results and certain practical value. At the end of the paper, the author briefly discussed the application of the smart city monitoring model in the reconstruction of old cities, hoping to provide certain guidance for further integration in the future.

Key words: Data-Driven, MFD construction, Multi-Source Data, Old city reconstruction, Urban intelligent monitoring

1. Introduction. Urban road networks are an essential component of urban infrastructure that directly affects people's everyday life. Urban road networks are under rising load strain due to the city's population and vehicle count, which frequently causes traffic jams and accidents that seriously impair urban development. Therefore, real-time perception and prediction of the operational status of urban road networks, as well as the identification of high-risk areas for traffic accidents in urban geographical space, are of paramount importance for enhancing the operational efficiency and traffic safety of urban road networks. This paper will take this as the research goal to solve the intelligent detection and recognition of urban road network.

The selection of traffic state evaluation indicators exhibits considerable variability across different scenarios. Recognizing the importance of data fusion between these diverse indicators, the concept of data fusion emerged in the late 20th century, driven by the rapid development of information technology [10]. Notably, Choi and Chung's [1] 2001 study on travel time using fusion coil sensor and GPS data laid the foundation for subsequent traffic flow data fusion. Yang Zhaosheng et al. [2] were pioneers in China, utilizing data fusion to meet information accuracy requirements in the ATMS subsystem of intelligent transportation systems. Henry[3] and others emphasized the real-time dynamics and data quality requirements for data fusion, playing a pivotal role in advancing the application of this technology in transportation. Xu Tao et al. [4] integrated urban road information, eliminating "information islands" and enhancing traffic flow parameters through Bayesian estimation fusion, ultimately identifying road states with fuzzy logic. Ding Yue et al. [5] proposed a multi-source relational data fusion framework, comprising pattern matching, entity alignment, and entity fusion, facilitating rapid pattern matching and providing a unified data view for analysis.Addressing the challenge of identifying traffic network conditions, common indicators include average speed, travel time, and travel delay. Ehrlich categorized traffic status indicators into time and distance, concluding that time-based indicators more accurately reflect travelers' perceptions [6]. Washburn used video data to calculate traffic flow density, evaluating road traffic conditions based on subjective traveler experiences [7]. Taylor et al. established a congestion coefficient model, incorporating GIS technology to build a congestion information system [8]. Kerner employed floating car data to establish a threshold model for evaluating congested sections [9].The average

*Eurasia Art and Design School, Xi'an Eurasia University, Xi'An, Shaanxi Province, 710065, China (Yi1Wang@outlook.com)

speed of a road section, known for simplicity and high reliability, stands out as one of the most commonly used indicators for traffic operation status evaluation [10,11]. Jiang Tao applied pattern recognition to classify traffic network states and predict future states based on current traffic conditions [12]. Sun Ya used data mining to extract new traffic status information from extensive data, achieving successful classification results through real-time traffic flow data collection [13]. Wang Meihong proposed a method to calculate regional traffic congestion correlation based on spatio-temporal association rules [14].Scholars have increasingly focused on the Macroscopic Fundamental Diagram (MFD) method for discriminating road network traffic operation status. Xu et al. derived optimal cumulative intervals and corresponding average traffic flow density states, categorizing traffic operation into free flow, optimal accumulation, and congestion [72]. Liu and Xu proposed using the standard deviation of the number of vehicles to supplement traffic flow or vehicle accumulation in traffic operation state division [73]. Daganzo and Gayah identified branch points in road networks, leading to multi-valued and unpredictable MFDs when exceeding certain densities, categorizing networks into stable and unstable states [74]. Haddad Geroliminis explained the generation of branch points, theoretically deducing equilibrium points under a dual system and proposing calculated and unstable boundaries [75]. Aboudolas et al. analyzed traffic ladder diagrams for urban road networks, differentiating between unsaturated, incomplete saturated, supersaturated, and deadlock states, recommending different control strategies for each [76]. With an increasing number of vehicles equipped with location devices, a considerable amount of vehicle Global Positioning System(GPS) data has emerged, leading to the gradual rise of research on road traffic state discrimination based on floating car GPS data. For example, based on massive GPS data analysis of residents' daily travel patterns, the regularities of road traffic operational status can be identified, including the recognition of urban hotspots for travel [20][21]. Lin [22] proposed a combination of spectral clustering techniques and branches based on MFD constructed from floating vehicle data Road network state identification method based on vector machine algorithm.

Urban road traffic is a complex network system; its state is challenging to summarize by a single parameter and often needs to combine a variety of traffic flow parameters and system methods to identify. However, the current research is usually carried out for a single parameter, and only some are carried out for multi-source data fusion. At the same time, noise processing of multi-source data is also a big challenge, and there are few scholars in this area of algorithm research. This paper mainly focuses on the research of multi-source data fusion and data noise reduction algorithm, and builds an intelligent monitoring model of urban road network based on this. In response to these issues, this paper proposes a data denoising model based on the wavelet algorithm and combines it with our proposed MFD multi-source data fusion construction, providing a theoretical foundation for building an intelligent urban traffic monitoring model. We then create a multi-source data-based urban traffic congestion monitoring model and a data-driven urban traffic safety environment monitoring model. Experimental results demonstrate that these models have good performance and practical value. In conclusion, we briefly discuss the application of the urban traffic intelligent monitoring model in the renovation of old city road networks, aiming to provide some guidance for future research.

2. Research on data processing in complex urban environment.

2.1. Research on data noise reduction model. In the field of traffic detection, more and more high-speed and high-precision sensors are being used due to the growing complexity of transportation networks. Traffic detector fault identification and prompt data rectification are critical in order to prevent sensor malfunctions that might jeopardize system safety and result in financial losses [23]. Taking loop detectors as an example, a single intersection may require the use of dozens or even more loop detectors to obtain more accurate traffic parameters such as traffic flow, vehicle speed, and occupancy. However, due to the variability of the natural environment and the inconsistency of the hardware and software attributes of front-end sensors, the rapid growth of data leads to a simultaneous increase in random noise and faulty data. Some fault information is masked by noise, and the distribution of faults becomes more multiscale. This renders traditional fault diagnosis methods based on analytical models inadequate to meet the safety requirements of current transportation systems. Therefore, how to use the accumulated large amount of offline traffic data, extract data features, analyze their intrinsic patterns, monitor online traffic data in real-time and effectively, and accomplish fault detection and diagnosis has become one of the hot topics of concern for many experts and scholars. Wavelet analysis is used to obtain the multiscale properties of so-called fault information, which might occur in many frequency bands.

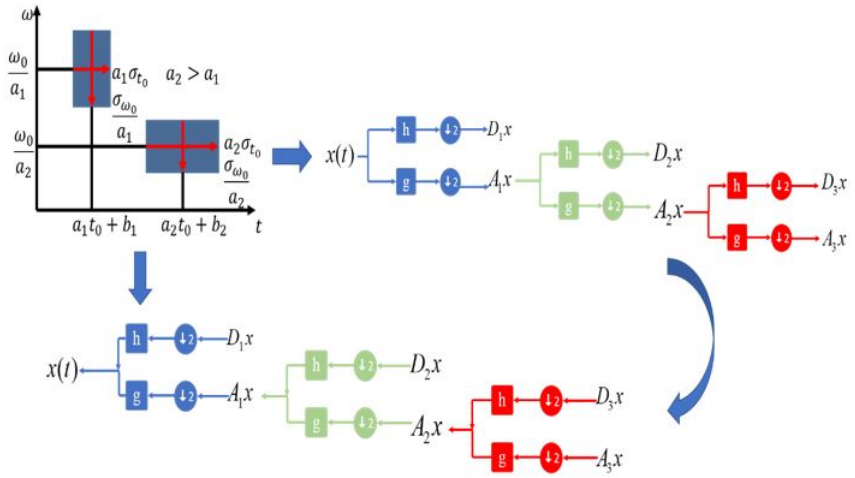

Fig. 2.1: Time-frequency coordinates of wavelet transform and Mallat algorithm

Let f(t) be a finite energy signal; the discrete wavelet transform of this signal can be defined as:

$$(W_\psi f)(a,b) = \langle f, \psi_{a,b} \rangle = |a_0|^{-m/2} \int_{-\infty}^{+\infty} f(t)\bar{\psi}(t)dt \tag{2.1}$$

Where $?_{a,b}$ is called the generating or basis function of the wavelet transform,

$$\psi_{a,b}(t) = a^{-1/2}\psi(\frac{x-b}{a}), a > 0, b \in R \tag{2.2}$$

The displacement factor is denoted by 'b' in equation (2.1), while the scale factor is represented by 'a'. As illustrated in Figure 1, a time-frequency coordinate system is established for wavelet transformation. In wavelet transformation, the position of the time window is solely influenced by the displacement factor. Therefore, as the scale factor increases, the time window widens, the frequency window narrows, and the center of the frequency window shifts towards the low-frequency direction. Conversely, with a narrower time window, the frequency window widens, and the center of the frequency window moves towards the high-frequency direction. The essence of wavelet transformation lies in manipulating these two factors to construct a combination that can represent any signal in space [24].By utilizing the scale factor, it is possible to perform a tower-like decomposition of a specific signal in space, as depicted in Figure 2.1, following the classical Mallat algorithm [25]. This algorithm provides a computational method for wavelet decomposition and reconstruction, simplifying the overall wavelet calculations.

There is an impulse response function in the Mallat algorithm: h(n). As a result, the following definitions for the wavelet and scale functions are given:

$$\begin{cases} \varphi(t) = \sum_n h(n)\varphi(2t-n) \\ \psi(t) = \sum_n g(n)\psi(2t-n) \end{cases} \tag{2.3}$$

In formula (2.3),$g(n) = (-1)^{1-n}h(1-n)$, The signal x(t) is decomposed by Mallat, and the scale is set to j(j≥1). Approximate signal and detailed signal obtained by decomposition are respectively:

$$\begin{cases} A_jx(t) = \langle x(t), \varphi_{j,k}(t) \rangle = 2^{-j/2} \int x(t)\varphi(2^{-j}t - 2k)dt \\ D_jx(t) = \langle x(t), \psi_{j,k}(t) \rangle = 2^{-j/2} \int x(t)\psi(2^{-j}t - 2k)dt \end{cases} \tag{2.4}$$

It can be found in equation (2.4) that the process of decomposing signal x(t) is the process of decomposing it step by step from scale j to j+1. That is, the process from low resolution to high resolution. It ultimately broke

down into two signals: an approximation signal, $D_j x$, and a detailed signal, $A_j x$, both at a high frequency [26]:

$$\begin{cases} A_{j+1}x = \sum_k h(k-2n)A_j x \\ D_{j+1}x = \sum_k g(k-2n)A_j x \end{cases} \quad j \geq 1 \tag{2.5}$$

Equation (2.6) is the Mallat wavelet reconstruction formula:

$$x = \sum_{\mathrm{k}=1}^{k} h(n-2k)A_{j+1} + \sum_{k=1}^{k} g(n-2k)D_{j+1} \quad j \geq 1 \tag{2.6}$$

After a finite energy signal undergoes wavelet transformation, it is decomposed into a set of detail signals and approximation signals. Each sample point of every signal has its wavelet decomposition coefficient $\omega_{j,k}$. When the signal contains noise, the noise is also decomposed along with the host signal. Therefore, theoretically, the noise may be removed if the wavelet decomposition coefficients of the noise are found and processed, and then all signals are submitted to wavelet reconstruction. The fundamental idea of wavelet threshold denoising is to target certain characteristics of noise in the signal, as well as the differences between noise and signal obtained after wavelet decomposition. A critical threshold is set, and the wavelet decomposition coefficients obtained after decomposition are compared with this threshold. If the result is less than the threshold, the coefficient is considered to belong to a normal signal and is retained without processing. This portion of wavelet coefficients is left untouched [27].Conversely, if the result is greater than the threshold, the coefficient is considered to be from noise and needs to be zeroed out or processed through a specific threshold function. This yields an estimate for this portion of wavelet coefficients to replace the original ones. Once all wavelet coefficients are processed, wavelet reconstruction is performed to achieve the denoising effect.The key to wavelet threshold denoising lies in finding an appropriate threshold function, also known as a threshold rule. Hard, soft, and semi-soft threshold functions are the three primary categories of conventional wavelet threshold functions [28]. This study suggests enhancements to the previously described semi-soft threshold denoising approach, based on substantial testing and empirical analysis:

Let the high frequency signal be $Wa_{j,k}$, then The formula for estimating the noise standard deviation is as follows:

$$\sigma_j = \frac{1}{0.6745} \times \frac{1}{N} \sum_{K=1}^{N} |Wa_{j,k}|, 1 \leq j \leq J \tag{2.7}$$

Since the real original signal's SNR varies, the threshold's setting must also be adjusted to reflect the current circumstances. The following is the unified threshold formula found in the body of available literature [29]:

$$\lambda_{1,j} = \sigma_j \sqrt{2\log(N)} \tag{2.8}$$

The high-frequency signal coefficients in the J group are produced once the signal has been broken down to the scale of J. Each group's wavelet coefficients are sorted in absolute value from small to big, yielding the following vector:

$$P = \left[\ Wa_{j,n} \ \right], 1 \leq n \leq N \tag{2.9}$$

The evaluation vector under the JTH wavelet coefficient is computed using this vector:$R = \left[\ r_n \ \right], 1 = \mathrm{n} = \mathrm{N}$ .Where:

$$r_n = \sum_{k=1}^{n} Wa_{j,n} + (N-i)Wa_{j,n} + (N-2n)\sigma_j^2 \tag{2.10}$$

After sorting the evaluation vector's interruption value from big to small and using the lowest value as the approximation error, the associated wavelet coefficient $Wa_{j,m}$ is discovered. The wavelet coefficient is used to determine the threshold value of the J-layer wavelet decomposition in the following manner:

$$\lambda_{a,j} = \sqrt{CD_{\min}} \tag{2.11}$$

The J-layer wavelet decomposition's threshold selection function is as follows:

$$\lambda_j = \begin{cases} \lambda_{1,j} & ,(P_{a,j} - \sigma_j^2 < \rho_{N,j}) \\ \min(\lambda_{1,j}, \lambda_{a,j}) & ,(P_{a,j} - \sigma_j^2 \geq \rho_{N,j}) \end{cases} \tag{2.12}$$

In this case, $\rho_{N,j}$ represents the wavelet coefficient vector's minimum energy level, and $P_{a,j}$ is the average value of the wavelet coefficient's absolute value. The following is the calculating formula:

$$P_{a,j} = \frac{1}{N} \sum_{k=1}^{N} W a_{j,k} \tag{2.13}$$

It is necessary to restore the signal to its original state since the calculated wavelet coefficient, or wavelet coefficient, is believed to be the result of noise. As a result, the original wavelet coefficient value is substituted for the estimated wavelet coefficient value through a series of computations, and wavelet reconstruction is ultimately used to accomplish the goal of noise reduction. The J-layer wavelet high-frequency signal's noise intensity is reflected by the introduction of a coefficient $\Gamma(\sigma_j)$ representing noise intensity. The following is the calculating formula:

$$\Gamma(\sigma_j) = \sqrt{\sigma_j / A_j} \tag{2.14}$$

In equation (2.14), A_j represents the amplitude of the high-frequency partial coefficient of the J-layer wavelet.

The calculation formula of wavelet coefficient estimate is given:

$$w_{j,k} = \begin{cases} w_{j,k} - \Gamma(\sigma_j) \times \lambda_j, & w_{j,k} > \lambda_j \\ w_{j,k} + \Gamma(\sigma_j) \times \lambda_j, & w_{j,k} < -\lambda_j \\ 0, & -\lambda_j \leq w_{j,k} \leq \lambda_j \end{cases} \tag{2.15}$$

Here are the specific actions to do in order to enhance the wavelet threshold denoising algorithm: The high-pass filter h and low-pass filter g are configured, and the original signal x(t) is discretized. The wavelet is disassembled by the J layer. The wavelet coefficients and the wavelet detail signal amplitude A_j of the layer decomposition are obtained. Based on the precise signal coefficients of each layer, the noise standard deviation σ_j and noise intensity coefficient $\Gamma(\sigma_j)$ of each layer are computed. Each layer signal's unified threshold, $\lambda_{1,j}$, is computed. The computation involves determining the adaptive threshold $\lambda_{a,j}$ for each layer signal, the lowest energy level $\rho_{N,j}$ of the layer's wavelet coefficients, and the average value $P_{a,j}$ of the absolute value of the layer's wavelet coefficients. Equation (2.12) is used to compute each layer's wavelet threshold. The wavelet coefficients are adjusted to complete the threshold denoising on this scale, and finally, the wavelet reconstruction is carried out according to the Figure 2.1.

Based on this as the main idea, simulation experiments are conducted on existing data. The results are shown in Figure 2.2. The studies demonstrate that, under the assumption of maintaining the majority of the fault information, the enhanced wavelet threshold denoising described in this study can suppress and eliminate noise better. Furthermore, the denoising findings are more appropriate for diagnosing data faults in the future.

2.2. MDF fusion construction for complex data scenarios. In the preceding section, a method model for data denoising was introduced. However, in complex data scenarios, it is often necessary to perform multi-source data fusion rather than just data denoising. This subsection will further discuss data processing based on this requirement. The construction of the MFD for road networks is the foundation of estimating traffic capacity and discerning the traffic operational status of road networks in this paper, based on the MFD method.

Addressing the potential issue of inconsistent MFD parameter estimation results in the process of constructing MFD based on multiple data sources, this paper proposes a framework and method for MFD fusion construction based on data reliability under the context of multi-source data, as illustrated in Figure 2.3.

The average traffic flow and average traffic flow density of the road network may be estimated using vehicle trajectory data and sectional detection traffic flow data. Several research have employed diverse data

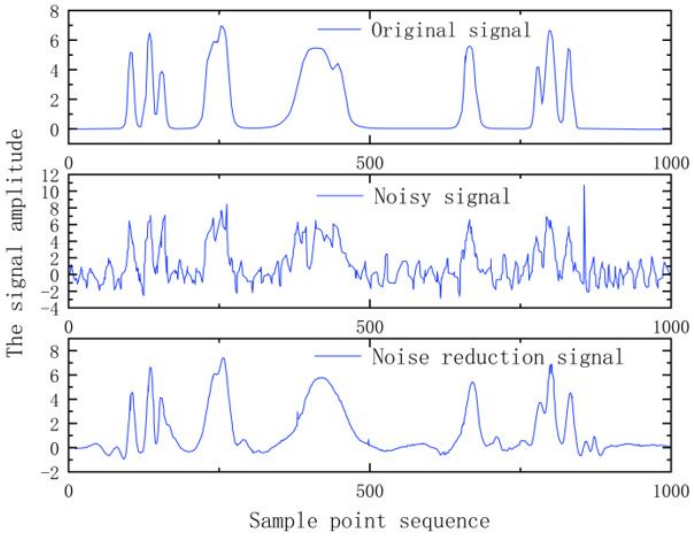

Fig. 2.2: Signal noise reduction effect diagram

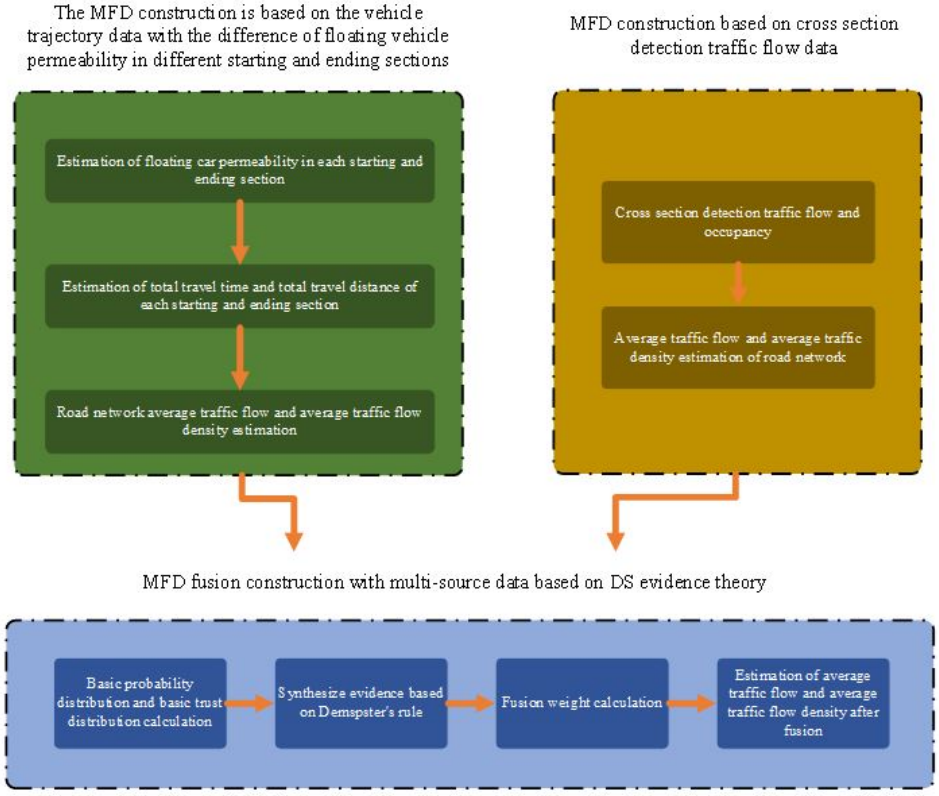

Fig. 2.3: MFD fusion construction framework based on multi-source data

sources to estimate different characteristics of MFD traffic operation states in order to get reliable estimates of the parameters of different traffic operation states in the MFD. The average traffic flow density and average traffic flow under various data sources are estimated using different assumptions, though. When estimating the average traffic flow density using vehicle trajectory data, it is common to assume that there are enough floating vehicle samples and that the penetration rates are distributed uniformly over space. On the other hand, the assumption that the traffic flow at the section represents the traffic flow of the road segment is typically used when estimating the average traffic flow based on sectional detection traffic flow data. As a result, outcomes of model parameter estimate under various data sources may differ. Equations (2.16) through (2.17) illustrate the fundamental foundation for the MFD fusion building in this study, taking the validity of data sources into consideration.

$$q(t) = \alpha(t) \times q^{\mathcal{F}}(t) + [1 - \alpha(t)] \times q^{\mathcal{M}}(t) \tag{2.16}$$

$$k(t) = \beta(t) \times k^{\mathcal{F}}(t) + [1 - \beta(t)] \times k^{\mathcal{M}}(t) \tag{2.17}$$

Among them, q(t) represents the average road network traffic flow estimated by multi-source data fusion during period t (vehs/h),$q^{\mathcal{F}}(t)$ represents the average road network traffic flow based on vehicle track data during period t (vehs/h),the average traffic density on a road network, measured in vehicles per kilometer for a certain time t is represented by the symbol $k^{\mathcal{F}}(t)$,$q^{\mathcal{M}}(t)$ represents the road network average traffic flow (vehs/h) based on the traffic flow data detected in section during the period t, $k^{\mathcal{M}}(t)$ represents the road network average traffic flow density (vehs/km) based on cross section detection traffic flow data during period t, the average traffic flow density (vehs/km) for the road network during period t, as determined via multi-source data fusion, is denoted by k(t). α(t) represents the floating vehicle data weight in the average road network traffic flow fusion estimation during period t; [1−α(t)] represents the road during period t The weight of traffic flow data detected by section in the fusion estimation of network average traffic flow,β(t) represents the weight of floating vehicle data in the fusion estimation of network average traffic flow density during the t period,[1−β(t)] represents the weight of traffic flow data detected by section in the fusion estimation of network average traffic flow density during the t period.

The process of figuring out the fusion weights of two different kinds of data sources is the basis of the fusion construction. This research aims to ascertain the fusion weights of diverse data sources based on the mean and variance distributions (reliability) of the average road network traffic flow density and average traffic flow across different periods of multiple days, utilizing the Dempster-Shafer(DS) evidence theory [30]. In order to illustrate the determination procedure, the following uses the floating vehicle data weight α(t) in the fusion estimate of road network traffic flow characteristics during t period as an example.

First, as indicated by Equation (2.18), the recognition framework of the DS evidence inference model ??(t) is built based on two types of data sources:

$$\Theta(t) = \{q^{M}(t), q^{\mathcal{F}}(t)\} \tag{2.18}$$

where $q^{\mathcal{M}}(t)$ and $q^{\mathcal{F}}(t)$ respectively represent the average traffic flow of the network in the t period based on the traffic flow data detected by the cross section and the vehicle track data.

Given that q∧M (t) and q∧F (t) originate from distinct data sources, it is possible to regard them as mutually exclusive. Equation (2.19) therefore displays the power set of all elements in

$$??(t) : \begin{cases} 2^{\theta(t)_0} = \{\emptyset, X_1(t), X_2(t), X_3(t)\} \\ X_1(t) = \{q^{M}(t)\} \\ X_2(t) = \{q^{\mathcal{F}}(t)\} \\ X_3(t) = \{q^{\mathcal{M}}(t) \cup q^{\mathcal{F}}(t)\} \end{cases} \tag{2.19}$$

Among them, $\emptyset$ as the empty set, lots the $\|\theta(t)\|\_0$ for ?? (t) L - 0 norm, $X_1(t)?X_2(t)?X_3(t)$ for ?? loophole (t) of the collection, $X_1(t)$ said the decision to cross section under the detection of traffic flow data of road network traffic flow (vehs/h) on average, $X_2(t)$means that the decision is the average traffic flow of the

road network under vehicle track data (vehs/h) $X_3(t)$ is an uncertain decision, indicating that it is impossible to distinguish which decision is $X_1(t)$ or $X_2(t)$. $q^{\mathcal{M}}(t)$ and $q^{\mathcal{F}}(t)$ have the same meaning as equation (2.18).

Second, two key ideas are included in DS evidence theory: The demspster evidence synthesis rule and fundamental trust distribution, which is also referred to as proof distribution or evidence function. A quantitative measure of the level of support and evidence for each choice is the basic trust distribution. The rule of evidence synthesis is a thorough examination that considers several pieces of data in relation to each choice. Make s:$2^{\|\theta(t)\|_0}\to[0,1]$, said s is a basic trust distribution $2^{\|\theta(t)\|_0}\to[0,1]$ mapping,it must meet the following requirements:

$$\begin{cases} s_i(\emptyset)=0 \\ \sum_{X_j(t)\subset 2^{\Theta(t)_0}} S_i\left(X_j(t)\right)=1 \end{cases} \qquad \mathrm{i}=1,2;\mathrm{j}=1,2,3 \tag{2.20}$$

where $S_i\left(X_j(t)\right)$ represents the degree of support for decision $X_j(t)$ by the evidence provided by the I-th data source, and its value is the basic trust assignment value of decision $X_j(t)$, the basic trust value of the empty set is 0, and the sum of the trust values of other subsets is 1. In this example, i stands for the evidence provided by the I-th data source, j for the JTH decision choice.

In this work, the basic probability allocation function of choice $p_i\left(X_j(t)\right)$ under each data source determines the fundamental trust allocation:

$$s_i\left(X_j(t)\right)=p_i\left(X_j(t)\right)/\sum_{j=1}^{3}p_i\left(X_j(t)\right),i=1,2 \tag{2.21}$$

In each data source, the fundamental probability distribution function of the decision $X_j(t)$ is represented by $p_i\left(X_j(t)\right)$.

Hypothesis $p_i\left(X_j(t)\right)$ meet the $\varphi_i(v)\sim N\left(\mu_i(t),\sigma_i^2(t)\right)$, the $_i(t)$ and$s_i^2(t)$ in the first class I average traffic flow data source under the historical data network $f_i(v)$ in t time mean and variance (?? = 1, 2), $p_i\left(X_j(t)\right)$ can be calculated according to equation (2.22).

$$\begin{cases} p_i(\bar{\emptyset})=0 \\ p_i(X_1(t))=\int_{\mu_1(t)-1/2}^{\mu_1(t)+1/2}\frac{1}{\sqrt{2\pi}\sigma_i(t)}exp\left\{-\frac{1}{2}\left(\frac{v-\mu_i(t)}{\sigma_i(t)}\right)^2\right\}dv \\ p_i(X_2(t))=\int_{\mu_2(t)-1/2}^{\mu_2(t)+1/2}\frac{1}{\sqrt{2\pi}\sigma_i(t)}exp\left\{-\frac{1}{2}\left(\frac{v-\mu_i(t)}{\sigma_i(t)}\right)^2\right\}dv \\ p_i(X_3(t))=\frac{1}{2\pi\sigma_l^2(t)}\left\{\begin{array}{l}\int_{\mu_1(t)-1/2}^{\mu_1(t)+1/2}exp\left\{-\frac{1}{2}\left(\frac{v-\mu_i(t)}{\sigma_i(t)}\right)^2\right\}dv \\ \times\int_{\mu_2(t)-1/2}^{\mu_2(t)+1/2}exp\left\{-\frac{1}{2}\left(\frac{v-\mu_i(t)}{\sigma_i(t)}\right)^2\right\}dv\end{array}\right\} \end{cases} \tag{2.22}$$

Based on the basic trust assignment $S_i\left(X_j(t)\right)$ of decision $X_j(t)$, the Demspster evidence synthesis process orthogonal and processed the basic probability assignment of multiple evidences. The evidence synthesis method of the two data sources is shown in Equation (2.23) :

$$s(X)=\begin{cases} K_D*\sum_{X_j(t)\cap X_j(t)=X}s_1\left(X_j(t)\right)s_2\left(X_j(t)\right) & X\neq\emptyset \\ 0 & X=\emptyset \end{cases} \tag{2.23}$$

Among them, j=1,2,3, $s(X)$,$X\subset 2^{??(\boldsymbol{t})?_0}$ is the synthetic trust distribution, $X_j(t)$ means that the decision is made under the evidence provided by the JTH data source, and the following formula is used to determine the conflict coefficient between the evidence offered by various data sources:

$$1/K_D=1-\sum_{x_j(t)\cap X_j(t)=\emptyset}s_1\left(X_j(t)\right)s_2\left(X_j(t)\right) \tag{2.24}$$

The closer $1/K_D$ is to 0, the greater the degree of conflict between $q^{\mathcal{M}}(t)$ and $q^{\mathcal{F}}(t)$ evidence provided by different data sources. When ,$1/K_D$=0, the sum under Dempster's synthesis rule does not exist.

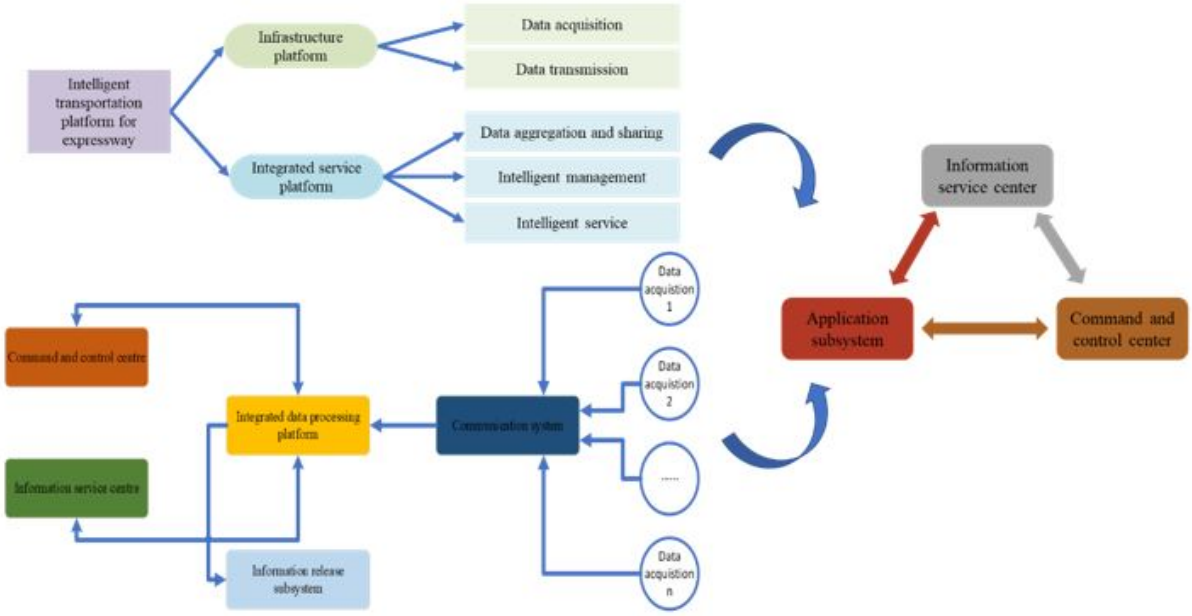

Fig. 3.1: ITS logical architecture diagram

The degree of disagreement between $q^{\mathcal{M}}(t)$ and $q^{\mathcal{F}}(t)$ evidence from various data sources increases with $1/K_D$'s proximity to 0. According to Dempster's synthesis rule, the total is null for $1/K_D$=0.

Equation (2.25) illustrates the approach used to calculate the $q^{\mathcal{F}}(t)$ weights. Lastly, fusion weights are generated based on synthetic trust allocation.

$$\alpha(t) = \frac{s(X_2(t))}{s(X_1(t)) + s(X_2(t))} \tag{2.25}$$

Assuming that α(t) is equivalent to equation (2.16), the synthetic trust allocation for each choice is represented by $S\,(X_1(t))$,$S\,(X_2(t))$ and $S\,(X_3(t))$, taking into account the evidence supplied by each data source.

3. Urban intelligent monitoring. With the accelerated process of urbanization and the development of the automotive industry, the contradiction between the existing urban road capacity and the continuously growing traffic demand has become increasingly acute, leading to a growing prominence of traffic congestion. Utilizing Intelligent Transportation Systems (ITS) for controlling and guiding traffic flow to alleviate congestion and provide a smooth and orderly traffic environment is an important means vigorously developed and applied by various countries. The so-called Intelligent Transportation System is established on the basis of improved road infrastructure, integrating new-generation information, the Internet of Things (IoT), and computer technologies into traffic management. It aims to create an information-real-time, accurate, widely-serviced, all-encompassing, highly efficient, and high-quality transportation and management system [31].In summary, ITS involves innovating traditional transportation systems through technology, creating a new type of transportation system that integrates informatization and intelligence [32]. The architecture of ITS is depicted in Figure 3.1. In this section, taking the complex urban road network as the research background, a combination of the two models mentioned earlier is performed. The goals of this include multi-source data-based urban traffic congestion monitoring and data-driven urban traffic safety environment monitoring.

3.1. Data-driven urban safety environment monitoring. A variety of factors are often considered while analyzing the traffic safety environment, such as the quantity (or frequency) of accidents, the nature of

the accidents (or types of vehicle collisions), and the severity of the accidents. The study of accident blackspots mostly uses accident frequency or quantity, with the usual objective being to pinpoint locations with a high accident rate. This kind of study often needs to pay more attention to including driver and vehicle conditions in the model and more attention to macroscopic traffic and road data. On the other hand, studies on accident types often concentrate on analyzing factors related to vehicles and roads, overlooking the impact of macroscopic traffic parameters and drivers. Therefore, a comprehensive analysis of accident severity that considers vehicles, drivers, macroscopic traffic, and road conditions can more comprehensively reflect the impact of accidents on the traffic safety environment. Thus, this paper selects accident severity as the research object.

Considering the comprehensiveness of data acquisition and the impact of the severity of injuries caused by accidents, a three-level classification method for accident severity is adopted, categorizing accidents into no-injury accidents, minor-injury accidents, and fatal or disabling accidents.

Various methods, including variable fusion, outlier detection and processing, are applied to restructure the relevant variables and data in the datasets. To ensure the timeliness and completeness of the data, this study uses accident and environmental data from recent years on Interstate 5 (I5) in Seattle, Washington. The data is collected from the National Highway Traffic Safety Administration's Fatality Analysis Reporting System (FARS) and the University of Washington's DRIVENet database. As the FARS dataset shares common labels such as TIME (time) and CASE NUMBER (accident number), the accident, vehicle, passenger, and environmental information in the information system table are fused based on these common labels and further combined with road information from the Washington University DRIVENet dataset, utilizing the MILEPOST (road mileage) label. The goal of restructuring the samples from multiple datasets is to make the samples suitable for model calculations. The initial complete sample labels include non-numeric symbols in textual expressions, mainly comprising some categorical variables. The restructuring of categorical variables involves label encoding and one-hot encoding.

On the basis of the above tests, in order to study the impact of different parameters on the traffic safety environment, the author introduced a factor importance index to evaluate the importance of different parameters. Considering that n-dimensional parameter vectors constitute the entire input space, the output vector Y corresponding to the first-stage partial derivative of the i_{th}-dimensional variable x_i can explain the sensitivity of this variable to the output. According to the chain rule of calculus, equation (3.1) is obtained:

$$\frac{\partial Y}{\partial X_i} = \frac{\partial Y}{\partial \alpha}\frac{\partial \alpha}{\partial X_i} \tag{3.1}$$

?? is the calculated value from the network hidden layer to the output layer, and equation (3.1) can be rewritten as:

$$\frac{\partial Y}{\partial X_i} = \sum_{j=1}^{L} W_{k,j}^{(2)}\frac{\partial H_j}{\partial X_i} g(\alpha)' \tag{3.2}$$

J refers to the JTH node of the hidden layer, and $\mathrm{W}_{k,j}^{(2.2)}$ refers to the weight from the hidden layer to the input layer. $g(a)^{'}$ is the activation equation from hidden layer to output layer.

Continuing to deduce according to the chain rule, Equations (3.3) and (3.4) can be obtained as follows.

$$\frac{\partial Y}{\partial x_i} = \sum_{j=1}^{L} W_{k,j}^{(2)}\frac{\partial H_j}{\partial \beta_j}\frac{\partial \beta_j}{\partial x_i} g(\alpha)^{'} \tag{3.3}$$

$$\frac{\partial Y}{\partial x_i} = \sum_{j=1}^{L} W_{j,i}^{(1)} W_{k,j}^{(2)} f(\beta)' g(\alpha)' \tag{3.4}$$

$\mathrm{W}_{j,i}^{(2.1)}$) refers to the connection weight from the input layer to the hidden layer, while $\mathrm{f}(\beta)^{'}$ is the activation equation from the input layer to the hidden layer.

Table 3.1: Influence factor importance and relative percentage

Factor	Categories	PSO optimization with w2		PSO optimization with w3	
		Factor Importance	relative percentage	Factor Importance	relative percentage
Type of accident	Non-road	0.108	77.3	0.080	63.4
Months	Non-road	0.088	63.1	0.102	81.0
Weekdays	Non-road	0.061	43.7	0.054	43.5
Type of location	Non-road	0.080	57.1	0.076	60.1
Road type	Non-road	0.078	55.6	0.063	50.2
Pavement	Road	0.052	37.4	0.056	44.3
Light	Non-road	0.069	49.1	0.071	56.1
Weather	Non-road	0.067	47.7	0.059	47.2
Road properties	Road	0.063	45.1	0.058	46.1
Driver gender	Non-road	0.024	17.4	0.017	13.1
Driver gear	Non-road	0.043	30.5	0.039	31.0
Vehicle type	Non-road	0.092	65.6	0.090	72.0
Curve direction	Road	0.032	23.2	0.028	22.1
Slope direction	Road	0.020	14.2	0.019	15.1
Driver age	Non-road	0.117	83.4	0.113	90.0
Car age	Non-road	0.140	100.0	0.126	100
Curve size	Road	0.047	33.5	0.045	35.4
Slope size	Road	0.056	40.1	0.052	41.3
Curve radius	Road	0.035	25.1	0.032	25.4
Curvature	Road	0.040	28.4	0.034	26.8

The relevance $R_i$of a single indicator can be calculated as follows (3.5) :

$$R_i = \sum_{j=1}^{L} W_{j,i}^{(1)} W_{k,j}^{(2)} / \sum_{i}^{N} \sum_{j=1}^{L} W_{j,i}^{(1)} W_{k,j}^{(2)} \tag{3.5}$$

In the above formula, all weights W can be obtained from the results of network simulation. In order to eliminate the randomness of the simulation, the simulation is performed k times, and the factor importance is obtained as follows (3.6) :

$$E(R_i) = \frac{1}{K} \sum_{k=1}^{K} R_i^k \tag{3.6}$$

By combining the calculation results of Particle Swarm Optimization(PSO) optimization network with the calculation formula of factor importance, the importance and relative percentage of factors can be obtained in Table 3.1 below.

It is evident from Table 3.1 and Figure 3.2 that the factor importance rankings derived from the two PSO optimization approaches are almost exact. Eight categories of factors—vehicle age, driver age, accident type, month, accident location type, road function, and lighting conditions—have a relative importance of more than 50%. Vehicle age and driver age are the two most important factors influencing accident severity, which aligns with both our intuition and experience. The impact of road factors is limited, with road type being the only one with a high importance ranking, indicating that specific types of accidents tend to occur on roads with specific functions. Combining the example of accident statistics mentioned earlier, it can be observed that the wet and snowy weather and mountainous roads in Washington State significantly increase the number and severity of accidents during the winter.

3.2. Urban congestion monitoring based on multi-source data. The variation of traffic states is regarded as random because of the high level of complexity and uncertainty associated with the spatiotemporal

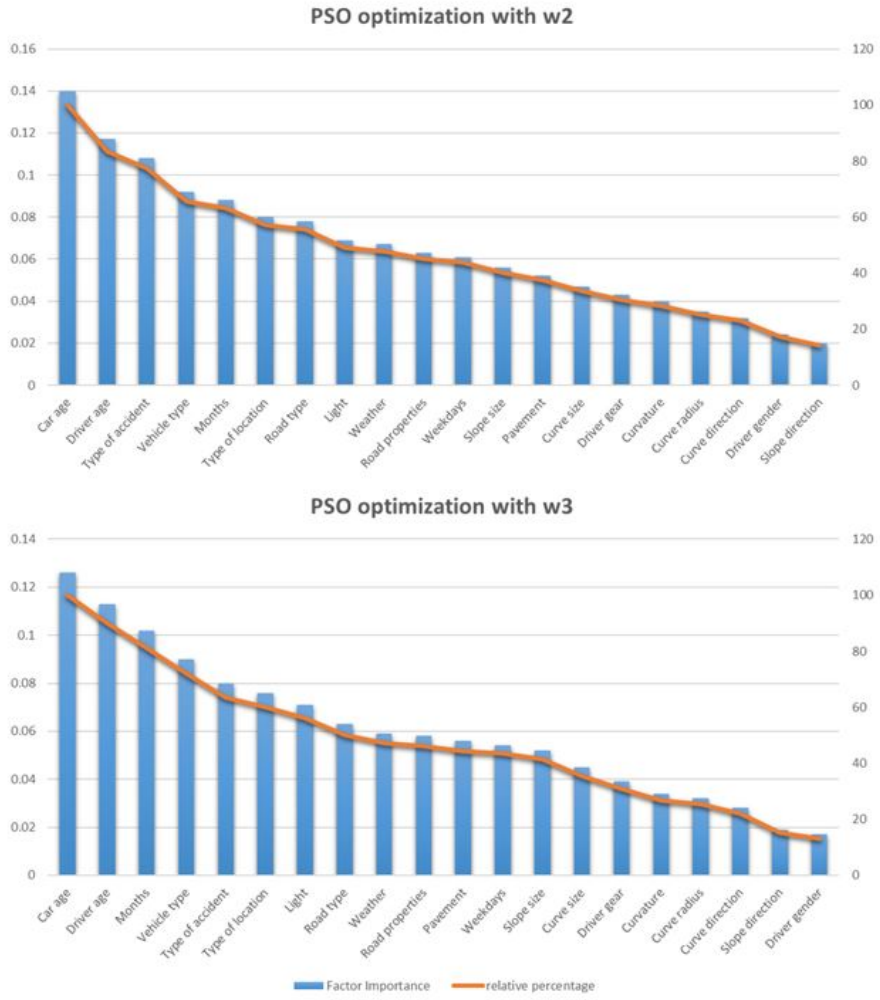

Fig. 3.2: Factor importance and relative percentage

state of traffic systems. To deeply explore the correlations between various traffic data, it is necessary to integrate and analyze multiple data sources to extract the hidden operational patterns of traffic states buried within the data. This chapter, based on various traffic flow parameter data, employs the MFD for data fusion. It uses a Genetic Algorithm-Fuzzy C-Means (GA-FCM) algorithm to classify road traffic states. Road traffic states can then be predicted by combining the anticipated outcomes of traffic flow parameters. The effectiveness of this method in reflecting road traffic conditions is confirmed by experimental results.

The Genetic Algorithm (GA) is a model parameter optimization algorithm that relies on natural selection and genetic mechanisms. FCM is a commonly used fuzzy clustering algorithm for data classification and cluster analysis. The steps of optimizing the FCM model using GA are as follows: 1) Determine the fitness function: The fitness function, which evaluates the quality of each individual, is crucial for GA. In the FCM model, the fitness function can be selected as an evaluation index for clustering effects, such as clustering accuracy or clustering entropy. 2) Determine the encoding method: GA needs to encode each individual into chromosomes for genetic operations. In the FCM model, each individual can be encoded into a set of fuzzy cluster centers, where each chromosome contains multiple genes representing cluster centers. 3) Initialize the population: Randomly generate an initial population, with each individual representing a random set of cluster

centers. 4) Selection operation: Based on the fitness function, select some excellent individuals as parents for the next generation. 5) Crossover operation: Perform crossover operations on parent individuals to generate new offspring individuals. In the FCM model, single-point crossover or multi-point crossover can be used. 6) Mutation operation: Introduce randomness to the offspring individuals through mutation operations to increase population diversity. In the FCM model, random perturbation or random replacement can be applied. 7) Evaluate fitness: Evaluate the fitness of the new generation, calculating the fitness value for each individual. 8) Repeat selection, crossover, mutation, and fitness evaluation operations until reaching the preset stopping conditions, such as reaching the maximum iteration times or convergence of fitness values. 9) Output the optimal solution: In the last generation of the population, select the individual with the highest fitness as the optimal solution, representing the optimal cluster centers.

Firstly, the GA-FCM clustering algorithm is applied to classify traffic flow parameters to obtain the optimal number of clusters and corresponding cluster centers for each state. Subsequently, a fuzzy clustering algorithm is utilized to partition the data, and based on the fuzzy cluster center results, different clusters representing traffic flow states are determined. These states are considered as prior knowledge for real-time traffic state identification. Finally, by calculating the membership degree of traffic flow data to each cluster center and the corresponding membership degree to different road traffic states, the traffic flow state at that moment can be determined. This process is accomplished using the membership degree function, selecting the traffic flow state with the maximum membership degree as the final identification result.

The selection range of the fuzzy factor is usually between 1 and 2. This parameter can suppress the influence of noise pollution by assigning larger weights to the membership function, reducing the impact of noise points on the FCM objective function during iterations. Starting from 1.0, experiments are conducted with a step size of 0.1, and the value of the fuzzy factor "m" is finally determined as 2. In the clustering analysis of traffic data, the number of clusters "C" plays a crucial role in the partitioning and identification of traffic states. To determine the optimal number of fuzzy partitions, the clustering validity function is commonly used as an evaluation criterion. This study employs the clustering validity function based on membership and squared membership weights to determine the number of clusters, as shown in Equation (3.7).

$$V_u = \frac{1}{2n}(\sum_{j=1}^{n}\sum_{i=1}^{c} u_{i,j}^2 + \sum_{j=1}^{n}\max_{i=1}^{c} u_i) \tag{3.7}$$

By computing when the number of clusters (C) is set to 4, the maximum fuzzy correlation value is achieved, indicating the highest effectiveness of data clustering. Therefore, in this experiment, we categorize the traffic states of road segments into four levels: smooth, moderately smooth, congested, and heavily congested. This categorization is based on the result obtained with a cluster number of 4.

The objective of the fuzzy mean clustering algorithm, optimized by a genetic algorithm in this study, is to identify a set of cluster centers and membership degrees that minimize the objective function. To achieve this goal, an iterative optimization approach is employed, iteratively updating the cluster centers and membership degrees to reduce the value of the objective function progressively. The objective function in this paper is the fuzzy mean clustering algorithm's loss function, as indicated by the following equation, because the underlying methodology uses the Fuzzy C-Means (FCM) algorithm for traffic state partitioning.

$$fit = minJ_f = min\sum_{j=1}^{c}\sum_{i=1}^{n}\left[u_j(d_i)\right]^m d_i - \nu_j^{\,2} \tag{3.8}$$

The selection of the number of iterations and the stopping error depends on the specific problem and dataset. Generally, a higher number of iterations leads to increased algorithm precision, but it also results in longer computation times. Therefore, it's advisable to choose an appropriate number of iterations based on the actual circumstances. The convergence of the algorithm can be observed to determine the number of iterations. If the algorithm has converged within a certain number of iterations, it can be stopped. The stopping error refers to halting iterations when the algorithm reaches a specific error range. The choice of the stopping error depends on the specific problem and dataset. If the dataset is noisy, a larger stopping error can be chosen to prevent overfitting.

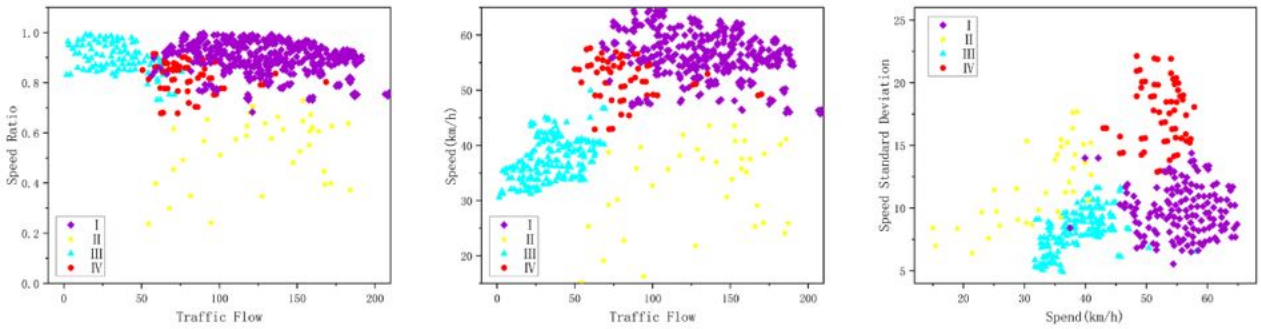

Fig. 3.3: Main Road Traffic Clustering Results

Conversely, if the dataset is clean, a smaller stopping error can be chosen to enhance precision. In practical applications, the number of iterations and stopping errors often need to be considered together. Typically, one can start with a small stopping error and then determine the number of iterations based on the convergence of the algorithm. If the algorithm has converged within a certain number of iterations, it can be stopped. If the algorithm hasn't converged, the number of iterations can be increased until convergence or reaching the maximum iteration limit. Considering these factors, this experiment sets the maximum number of iterations to 100 and the iteration-stopping error threshold to 1e-5.

This study intends to conduct an instance analysis using two different levels of road segment data. Clustering analysis will be performed using historical traffic flow parameter data to obtain a road traffic state discrimination method based on multi-source data. The fuzzy partitioning process of road traffic states based on GA-FCM includes the following steps: extracting features such as flow, speed, and occupancy rate from the original data, constructing the objective function, utilizing genetic algorithms to find the optimal clustering centers, and finally outputting the clustering results of traffic states.

According to the above parameter Settings, the clustering centers of four traffic states of the main road obtained by GA-FCM clustering algorithm are represented by $V=\{\nu_1,\nu_2,\nu_3,\nu_4\}^{\mathrm{T}}$, and the results of equation (3.9) and Figure 6 are obtained respectively.

$$V1=\begin{bmatrix} 31.82 & 55.88 & 0.04721 \\ 85.17 & 49.74 & 0.08338 \\ 148.7 & 50.37 & 0.2233 \\ 109.9 & 18.3 & 0.1975 \end{bmatrix} \tag{3.9}$$

In this context, the three columns of the clustering centers represent flow, speed, and occupancy, respectively. Each row corresponds to a specific traffic state, namely, smooth-flowing, moderately smooth-flowing, congested, and heavily congested.

From the graph, it can be observed that there is a certain continuity in the value ranges between different traffic states. Under different traffic states, certain parameter ranges exhibit overlapping characteristics. As the traffic state transitions from smooth-flowing to congested, the speed gradually decreases while the flow and occupancy increase. In a heavily congested state, where vehicles move at a slower pace, the flow gradually decreases, aligning with the fundamental principles of traffic operation.

V2 represents the clustering centers for the four traffic states on the expressway, and the partition results are illustrated in Equation (3.10) and Figure 3.4.

$$V2=\begin{bmatrix} 25.41 & 50.3 & 0.04761 \\ 71.95 & 44.45 & 0.08646 \\ 119 & 45.5 & 0.223 \\ 87.87 & 15.91 & 0.2007 \end{bmatrix} \tag{3.10}$$

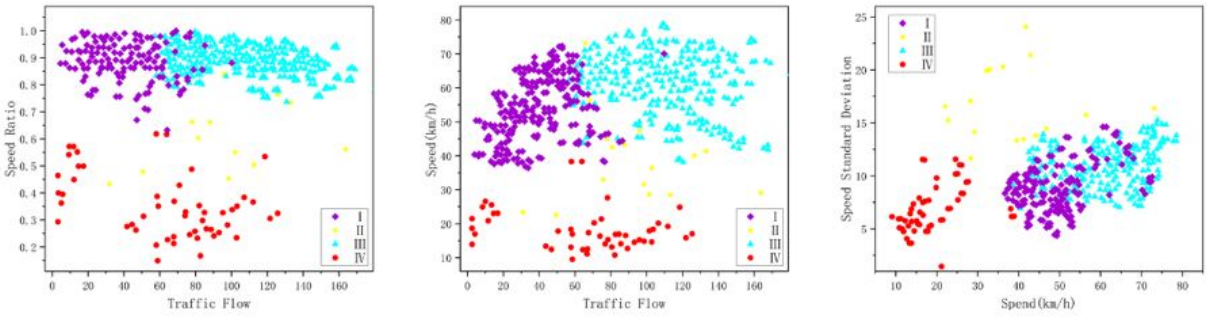

Fig. 3.4: Expressway Traffic Clustering Results

From the above figure, it is evident that, compared to arterial roads, expressways exhibit slightly higher traffic flow and speed. Expressways demonstrate relatively high traffic flow stability, with minimal speed differences among vehicles, resulting in a smoother traffic flow.

Different types of roads exhibit distinct characteristics in traffic flow parameters, necessitating reasonable planning and design to ensure the efficient operation and safety of the urban road network. Expressways typically have higher traffic volumes than arterial roads, as they handle a large number of passenger and freight vehicles. Arterial roads follow in terms of traffic volume, primarily accommodating internal traffic flows within the city, including commuting, commercial, and service-related traffic. Regarding speed, expressways usually have higher speeds than arterial roads due to their higher design speeds, wider lane widths, and absence of traffic signals, allowing vehicles to travel at higher speeds. Arterial roads have slightly lower speeds as traffic signals typically control them, and vehicles need to adhere to traffic rules and signal indications, resulting in relatively slower speeds.

Compared to the traditional FCM clustering algorithm, the GA-FCM clustering algorithm optimized by a genetic algorithm can more rapidly identify the optimal clustering partition. Additionally, both algorithms yield the same final value for the objective function, with the FCM algorithm consistently having a higher final value than the GA-FCM algorithm. The graph also illustrates that GA-FCM exhibits faster convergence and better stability compared to the traditional FCM clustering algorithm.

3.3. Research on old city reconstruction. Urban traffic intelligent monitoring provides unprecedented opportunities for the transformation of urban road networks. With continuous technological advancements, traffic managers can utilize advanced monitoring technologies and data analysis tools to achieve more efficient and intelligent road network designs. This process not only improves traffic flow but also contributes to enhancing the quality of life for urban residents.

Firstly, intelligent monitoring systems provide cities with comprehensive traffic data. By deploying cameras, sensors, and other monitoring devices, traffic managers can obtain real-time information on vehicle flow, speed, congestion, and more. These data form the basis for a deep understanding of the city's traffic conditions. Leveraging advanced artificial intelligence algorithms, managers can extract valuable information from massive datasets and identify traffic patterns, peak periods, and potential bottleneck locations. With this information, cities can engage in more precise traffic planning and road network improvements.

The data analysis capabilities of intelligent monitoring systems empower traffic managers to identify bottlenecks and congestion points, allowing targeted optimization of intersections and adjustments to traffic signals. Through real-time traffic flow control, congestion can be effectively alleviated, and road operational efficiency improved. This not only helps relieve traffic pressure but also reduces carbon emissions, enhancing urban air quality.

Real-time traffic information dissemination systems extend the capabilities of intelligent monitoring systems by providing drivers and citizens with real-time information through mobile apps, digital signage, social media

platforms, and more. This enables citizens to flexibly plan travel routes and choose alternative roads to avoid congestion. Additionally, it enhances the traffic awareness of participants, reducing the incidence of traffic accidents.

Data sharing and cross-departmental cooperation are crucial for achieving intelligent traffic management. Establishing data-sharing mechanisms among different traffic systems facilitates the collaborative operation of traffic systems. Through big data analysis, urban planners can better understand the city's traffic demands, providing a scientific basis for future road network transformations.

In addition to improving traffic efficiency, urban traffic intelligence monitoring should also focus on environmental sustainability. Combining traffic monitoring systems with environmental monitoring technologies allows cities to assess the impact of traffic on air quality and noise levels. Based on these assessments, corresponding measures can be taken, such as the establishment of green belts and the construction of sound barriers, to improve the living environment for urban residents.

In conclusion, road network transformation based on urban traffic intelligent monitoring represents a revolutionary attempt in urban traffic management. By fully leveraging advanced monitoring technologies and data analysis tools, cities can achieve a more intelligent, efficient, and environmentally friendly traffic system. This not only concerns the development of urban traffic but also affects the travel experience and quality of life for urban residents.

4. Conclusion. This paper makes a detailed discussion of urban intelligent monitoring, puts forward some algorithm models of intelligent city monitoring based on data-driven, and briefly discusses its application in the reconstruction and renewal of old cities. At the same time, it also puts forward new solutions for data noise reduction and multi-source data fusion in complex urban environments. Summarizing the relevant research in this paper, the following conclusions are drawn:

1. This paper proposes a data-driven signal denoising model using an improved wavelet threshold method. Three key points of wavelet threshold denoising are modified, and the detailed calculation formula and calculation flow of the denoising algorithm are given. Using the experimental data containing noise and fault to debug, the signal-to-noise ratio decreases significantly after noise reduction, and most types of fault information can still exist on multiple scales. Experiments show that the model can suppress and remove the noise well while retaining most fault information. The noise reduction results are more suitable for the late diagnosis of data faults and can be applied to constructing an urban intelligent monitoring model.
2. An MFD fusion structure is proposed to aim at the problem of multi-source data in urban traffic intelligence monitoring. An urban congestion monitoring model is constructed by combining the FCM model optimized based on a genetic algorithm. Compared with the traditional FCM clustering algorithm, the GA-FCM clustering algorithm optimized by the genetic algorithm can find the optimal cluster classification faster. In addition, the final value of the objective function of both algorithms is the same, and the final value of the objective function of the FCM algorithm is always higher than that of the GA-FCM algorithm.
3. The application of the monitoring model in the smart city in the reconstruction and renewal of the old city is discussed, with clear objectives, reasonable means and correct direction, which has certain guiding significance.

The research in this paper is a preliminary exploration of road network operation status and traffic capacity estimation after noise reduction of multi-source data using a wavelet algorithm. With the deepening of the research, the author feels the challenge of the research and still doubts the further application and development of the model in the reconstruction of old cities. Based on the current research, future research can be carried out from the following three aspects.

In the process of MFD fusion construction based on multi-source traffic perception data, limited by the acquisition of actual data, this paper only selected a small range of actual road networks to verify the method. In the future, when more realistic data are collected, the accuracy of the proposed method in the actual large-scale road network MFD fusion construction needs to be further verified. In the link of road network dynamic capacity estimation, this paper adopts the data-driven method, the premise of which requires the road network to have a relatively complete traffic operation state (unsaturated, saturated, and supersaturated). Therefore,

its application may be limited to the capacity analysis of road networks such as central urban areas. In the future, theoretical methods based on analytical modeling can be further considered. A road network capacity estimation method driven by data and model is constructed. The application of this model in the reconstruction of the old city is only a superficial discussion in this paper, without further analysis and calculation, which can be studied in the future.

REFERENCES

[1] Gao, X. & Wang, Y. Data fusion technology review. *Journal Of Computer Automatic Measurement And Control.*, 706-709 (2002)

[2] Yang, Z., Wang, S. & Ma, D. Review of basic traffic information fusion methods. *Highway Transportation Science And Technology.*, 111-116 (2006)

[3] El Faouzi, N., Leung, H. & Kurian, A. Data fusion in intelligent transportation systems: Progress and challenges–A survey. *Information Fusion.* **12**, 4-10 (2011)

[4] Xu, T., Yang, X. & Xu, A. Research on Data Fusion for Urban Road Traffic State Estimation. *Computer Engineering And Applications.* **47**, 218-221 (2011)

[5] Ding, Y., Wang, J. & Lu, W. Fusion of multi-source relational data. *Science In China: Information Science.* **50**, 649-661 (2019)

[6] D'Abadie, R. & Ehrlich, T. Contrasting Time-Based and Distance-Based Measures for Quantifying Traffic Congestion Levels: Analysis of New Jersey Counties. (2002)

[7] Washburn, S. & Kirschner, D. Rural Freeway Level of Service Based on Traveler Perception. *Transportation Research Record Journal Of The Transportation Research Board.* **1988** pp. 31-37 (2006)

[8] Taylor, M., Woolley, J. & Zito, R. Integration of the global positioning system and geographical information systems for traffic congestion studies. *Transportation Research Part C: Emerging Technologies.* **8**, 257-285 (2000)

[9] Kerner, B., Demir, C. & Herrtwich, R. Traffic state detection with floating car data in road networks. *Proceedings. 2005 IEEE Intelligent Transportation Systems.* pp. 44-49 (2005)

[10] Feng, X., Zhou, C. & Rong, J. Research on frequent traffic congestion on urban expressway based on speed characteristics. *Traffic Information And Safety.* **32**, 29-33 (2014)

[11] Sun, C., Zhang, H. & Chen, X. Road traffic operation evaluation based on multi-source floating vehicle data fusion. *Journal Of Tongji University (Natural Science).* **46**, 46-52 (2018)

[12] Ren, J., Ou, X. & Zhang, Y. Research on traffic status pattern recognition. *Highway Transportation Science And Technology.*, 63-67 (2003)

[13] Sun, Y., Qian, H. & Ye, L. Application of data mining algorithm in traffic state quantification and recognition. *Journal Of Computer Applications.*, 738-741 (2008)

[14] Wang, M., Xie, D. & Zhao, X. A method for calculating regional traffic congestion correlation based on spatiotemporal association rules. (2013,12,24)

[15] Xu, F., He, Z. & Sha, Z. Traffic state evaluation based on macroscopic fundamental diagram of urban road network. *Procedia-Social And Behavioral Sciences.* **96** pp. 480-489 (2013)

[16] Shuqing, L. & Jianmin, X. Urban traffic state analysis based on the macroscopic fundamental diagrams of the variability of vehicle densities. *2016 12th World Congress On Intelligent Control And Automation (WCICA).* pp. 1010-1015 (2016)

[17] Daganzo, C., Gayah, V. & Gonzales, E. Macroscopic relations of urban traffic variables: Bifurcations, multivaluedness and instability. *Transportation Research Part B: Methodological.* **45**, 278-288 (2011)

[18] Haddad, J. & Geroliminis, N. On the stability of traffic perimeter control in two-region urban cities. *Transportation Research Part B: Methodological.* **46**, 1159-1176 (2012)

[19] Aboudolas, K., Papageorgiou, M. & Kouvelas, A. A rolling-horizon quadratic-programming approach to the signal control problem in large-scale congested urban road networks. *Transportation Research Part C: Emerging Technologies.* **18**, 680-694 (2010)

[20] Qu, Z., Wang, X. & Song, X. Based on large data of urban taxi GPS hot travel road recognition method. *Journal Of Transportation Systems Engineering And Information Technology.* **12**, 238-246 (2019)

[21] Xu, X., Dou, W. & Zhang, X. A traffic hotline discovery method over cloud of things using big taxi GPS data. *Software.* **47**, 361-377 (2017)

[22] Lin, X. A road network traffic state identification method based on macroscopic fundamental diagram and spectral clustering and support vector machine. *Mathematical Problems In Engineering.* **2019** pp. 1-10 (2019)

[23] Xu, Z., Mo, X. & Xu, Z. Research review of road traffic detectors and their optimal layout methods. *Journal Of South China University Of Technology (Natural Science Edition).* **51**, 68-88 (2023)

[24] Xu, B., Cen, K. & Huang, J. A review of graph convolutional neural networks. *Chinese Journal Of Computers.* **43**, 755-780 (2020)

[25] Luo, D., Li, Y. & Luo, Z. Detection and analysis of hanging basket wire rope broken strands based on Mallat algorithm. *International Conference On Neural Computing For Advanced Applications.* pp. 518-532 (2023)

[26] Li, H., Zhou, Y. & Tian, F. A new adaptive wavelet thresholding function vibration signal denoising algorithm. *Journal Of Instruments And Meters.* **4**, 2200-2206 (2015)

[27] Chang, G. Adaptive wavelet thresholding for image denoising and compression. *IEEE Transactions On Image Processing.* **9** (2000)

[28] Guo, H., Jing, X. & Shang, Y. Research on vehicle license plate location based on wavelet transform and mathematical morphology. *Computer Technology And Development*. **20**, 13-16 (2010)
[29] Hou, P., Zhao, J. & Liu, M. License plate location method based on wavelet transform and line scan. *Journal Of System Simulation.*, 811-813 (2006)
[30] Qiu, D., Li, X. & Xue, Y. Analysis and prediction of rockburst intensity using improved DS evidence theory based on multiple machine learning algorithms. *Tunnelling And Underground Space Technology*. **140** pp. 105331 (2023)
[31] Practice Practice of comprehensive urban traffic management in Zhejiang Province under the background of traffic congestion control. *Urban Transportation*. **18**, 11-12 (2020)
[32] Xiao-Min, Z. Current situation analysis and suggestions on the development of intelligent transportation. *Journal Of Electronic World.*, 96-97 (2020)

Edited by: Zhengyi Chai
Special issue on: Data-Driven Optimization Algorithms for Sustainable and Smart City
Received: Jan 15, 2024
Accepted: Apr 28, 2024