An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
IJOCTA, Vol.16, No.4, pp.1560-1585 (2026)
https://doi.org/10.36922/IJOCTA026180071

RESEARCH ARTICLE

Organizational intelligence for real-time traffic estimation with real Taif sensor deployment

Mahmoud Rokaya[1*], Dalia I. Hemdan[2], Ashraf Alyanbaawi[3], Abdulqader M. Almars[4], Ghada Elmarhomy[5], and El-Sayed Atlam[6,7]

[1]Department of Information Technology, College of Computers and Information Technology, Taif University, Taif, Mecca Province, Saudi Arabia
[2]Department of Food Science and Nutrition, Faculty of Science, Taif University, Taif, Mecca Province, Saudi Arabia
[3]Department of Networks and Information Systems, College of Computer Science and Engineering, Taibah University, Yanbu, Medina Province, Saudi Arabia
[4]Department of Computer Science, College of Science, Northern Border University, Arar, Arar Province, Saudi Arabia
[5]Department of Information Systems, College of Computer Science and Engineering, Taibah University, Yanbu, Medina Province, Saudi Arabia
[6]Department of Computer Science, College of Computer Science and Engineering, Taibah University, Yanbu, Medina Province, Saudi Arabia
[7]Department of Computer Science, Faculty of Computer and Information, Tanta University, Tanta, Gharbiya, Egypt
mahmoudrokaya@tu.edu.sa, dalia.m@tu.edu.sa, ayanbaawi@taibahu.edu.sa, aalmars@nbu.edu.sa, gmarhomy@taibahu.edu.sa, elsayed.atlam@science.tanta.edu.eg

ARTICLE INFO

Article History:
Received: April 29, 2026
Revised: June 6, 2026
Accepted: June 8, 2026
Published Online: July 8, 2026

Keywords:
Traffic flow optimization
Intelligent transportation systems
Distributed optimization
Real-time traffic estimation
Graph-based learning
Multi-agent systems
Urban mobility management
Partial observability

ABSTRACT

Real-time urban traffic estimation and prediction under sparse sensing, partial observability, and varying demand persist as difficult optimization issues in intelligent transportation systems. Recent centralized deep models typically require large historical datasets, incur high computational costs, and have insufficient adaptability to dynamic environments. To this end, we explored the design of a scalable and robust framework for improving traffic reconstruction and forecasting with incomplete observations. To accomplish this goal, we presented an organizational intelligence model, in which the traffic network was modeled as a dynamic graph of lightweight local computational units assigned to each road segment or intersection. These entities work together in distributed optimization processes that include adaptive influence weighting, role specialization, and structural coordination. This architecture separates local computation from global coordination, enabling efficient training without increasing model depth. We tested the framework on benchmark traffic data and realistic urban environments in Taif, Saudi Arabia. The proposed model on the benchmark dataset reduced the test root mean-squared error from 0.5578 to 0.3530 (36.7%), and the test mean absolute error (MAE) from 0.3013 to 0.1614 (46.4%) relative to a centralized baseline. The validation loss decreased by 48.0%. On real-scale deployment units, total MAE fell from 6.6473 to 6.2060, while hidden-node MAE decreased from 6.8816 to 6.4787. Other experiments demonstrated lower inference latency, increased robustness in limited-data scenarios, and stable transfer to unseen environments. These results show that organizational intelligence is an efficient, scalable, optimization-oriented alternative for next-generation real-time urban traffic management.

*Corresponding Author

1. Introduction

Urban traffic systems are dynamic, networked, and partially visible environments that need to be estimated, predicted, and managed within operational constraints. Recent advances in distributed artificial intelligence (AI) have highlighted decentralized learning[1] and optimization in large-scale infrastructure, particularly in communication and transportation systems. Meanwhile, collaborative prediction, anomaly detection, and adaptive control for smart cities are widely implemented using multi-agent approaches.[2] Emerging paradigms, such as digital twins and metaverse-integrated traffic systems, extend beyond these to incorporate real-time perception, reasoning, and resource scheduling across urban infrastructures.[3]

All of these trends indicate an increasing focus on smart, optimization-guided transportation management. Extending these insights, many AI-based traffic prediction and control models have been developed to tackle traffic flow estimation and network optimization.[4] Most traditional solutions are based on centralized architectures and may not be scalable or provide real-time responsiveness. Comparison studies between centralized and distributed paradigms have found that distributed systems tend to be much more robust and dynamically flexible computational solutions[5] than centralized strategies in dynamic environments. Recent works have also contributed to improved collaborative decision-making for connected and autonomous vehicles through graph-based and multimodal solutions,[6] and online optimization and spatiotemporal decision-making approaches have been proposed for traffic conditions that change over time and exhibit heterogeneity.[7]

The emergence of multi-agent reinforcement learning (MARL) as an efficient framework for decentralized traffic optimization and control furthers adaptability.[2,8–10] Similar approaches evolve cooperative policies under repeated interaction with the environment. Multi-agent models[11] have been introduced to account for heterogeneous driving behaviors and their effects on network dynamics. In this regard, distributed MARL frameworks have proven to be effective for motorway control and large-scale traffic flow management,[12] whereas recent surveys document the growing importance of multi-agent systems and edge computing for scalable intelligent transport systems.[13] Simultaneously, spatiotemporal learning methods, such as graph neural networks and transformer architectures, have achieved strong predictive performance in traffic flow estimation. These techniques utilize graph structures and attention mechanisms to model spatial dependency and temporal development.[14–16] Decoupling patterns over the long term from short-term time has improved predictive accuracy.[17]

General research has supported the significance of graph neural networks in transportation analytics in large-scale studies,[18] and machine learning has also been adapted to maximize traffic signals and intersection control.[19] Recent hybrid-transformer technologies, memory-augmented networks, and higher-performance applications have also been applied to scalable model prediction and its quality.[20–23] Transportation systems are more than just prediction; optimization-oriented decision-support mechanisms that turn model outputs into operating strategies are further needs of those in transportation systems, thus moving beyond prediction. Early decision support systems (DSSs) frameworks laid the groundwork for group decision-making and transportation planning.[24–26]

These systems subsequently also included geographic knowledge systems, driver behavior modeling, and AI-based sustainable mobility planning systems.[27–29] Data-driven DSSs also greatly enhance the ability of large transportation decision-making.[30] However, a central optimization challenge remains: real traffic systems can only be observed at a partial level, meaning a limited set of sensors is available at each time instant. Uncertainty-aware learning, historical inference, and incomplete-observation modeling[31] have all been introduced to solve this problem. This has also been applied to partially observable environments in the background via multimodality and historical context.[32,33]

Further research on traffic sensor networks also underscores the importance of spatial correlation for observability and resilience.[34] Recent ontology-augmented and uncertainty-sensitive approaches now focus on robust optimization in fluctuating, uncertain contexts.[35–37] The Internet-of-Things (IoT)-based adaptive traffic management systems play a similar role in supporting the distributed and real-time optimization of urban mobility systems.[38] Complementary works involve computational experimentation frameworks, communication-based coordination approaches, and parallel control paradigms. The integration of big-data transportation analytics and DSS-based computational experiments has further enhanced the system-level optimization capability.[39] Real-time traffic control methods, such as communication-based approaches like vehicular ad hoc networks (VANETs), have been proposed for distributed coordination.[40] Parallel control and management paradigms incorporate computation, experimentation, and optimization in intelligent transportation systems.[41]

Recently, adaptive filtering and state estimation methods for congestion prediction and dynamic control techniques have been introduced for congestion forecasting and state estimation,[42] and long-term research in this field has been developed in the context of distributed intelligent transportation systems, with a main focus on distributed-based intelligent transportation systems.[43] However, the majority of conventional approaches are optimized for a small part of a single problem. Deep learning models prioritize accuracy, multi-agent systems emphasize decentralization, while the design and planning frameworks of DSSs focus more on "top-level" planning. An integrated model that jointly optimizes prediction, coordination, and the adaptive system structure is still missing. This limitation, especially under partial observability, changing demand trends, and real-time computing requirements, is important. Inspired by this deficiency, this paper proposes a new concept for organizational intelligence that reinterprets traffic estimation and prediction as a distributed optimization problem.

The traffic network is constructed as a small, cooperating computational model that operates only under local constraints regarding information and interacts flexibly and autonomously with neighboring nodes. The global performance emerges as a performance-oriented coordination, role specialization, and structural adaptation process. Instead of deepening the model, the design maximizes the organization of computation throughout the network. The main contribution of this study is to link intelligent transportation systems with a modern optimization perspective through an organizational lens. In contrast to fixed centralized architectures or pre-defined coordination rules, the proposed system constantly reorganizes itself to improve accuracy, scalability, and robustness during its estimation process. The framework tests the approach under realistic traffic conditions, including sparse sensing, limited training data, and dynamic conditions. From these results, organizational intelligence is positioned as an interesting optimization-oriented base level of the next-generation urban traffic management system.

The main contributions of this work are summarized as follows:

- Proposes a new organizational intelligence framework that models urban traffic systems as distributed, cooperative computational entities, shifting from model-centric to system-level optimization.
- Develops a distributed optimization approach for accurate traffic reconstruction and prediction from sparse, incomplete sensor data, addressing real-world observability limitations.
- Introduces a dynamic graph-based structure in which each road segment/intersection serves as a local computational unit, enabling scalable, efficient computation without increasing model complexity.
- Incorporates adaptive influence weighting, role specialization, and structural coordination, allowing the system to self-organize and respond to dynamic traffic conditions.
- Designs a framework that decouples local processing from global optimization, improving training efficiency and reducing computational overhead compared to centralized deep learning models.
- Demonstrates effectiveness in Taif, Saudi Arabia, with reduced MAE across both observed and hidden nodes, confirming its applicability in real urban deployments.
- Shows lower inference latency, improved robustness under limited-data scenarios, and stable performance when transferred to unseen locations.
- Establishes the proposed approach as a scalable, robust, and optimization-oriented alternative for real-time intelligent transportation systems.

2. Literature review

Intelligent transportation systems research has developed through numerous interconnected streams, including DSSs, distributed and multi-agent intelligence, spatiotemporal learning models, and uncertainty-aware traffic analysis. These streams, altogether, provide the underlying basis for contemporary traffic prediction and estimation systems, but are still predominantly siloed in terms of their integration from decision-making to prediction to adaptive response in a single schema. DSSs and intelligent decision technologies in transport play a foundational role in research. Early work introduced structured systems of group decision-making and traffic management based on theoretical models of the role of computational tools for human operators and policymakers.[24–26] In addition, Geographic Information Systems (GIS), driver behavior modeling, and AI-based decision support have been used to further improve sustainable urban mobility.[27–29] The advent of big data analytics further facilitated data-driven decision-making, as DSS frameworks leveraged it to process transportation data and generate actionable insights at scale.[30,39] Complementary advancements include communication-based coordination methods, such as VANETs,[40] and parallel control and management models that combine computation, experimentation, and control in intelligent transportation approaches.[41,43] Recent work has used adaptive filtering or state estimation methods to enhance congestion prediction and adaptive decision-making in an ever-changing scenario.[42] However, DSS frameworks tend to be at a high level and depend on the outputs of underlying predictive models rather than directly integrating decision-making into the computational framework. Recently, distributed intelligence and multi-agent systems have become the paradigm to address large-scale traffic conditions. New developments in distributed AI emphasize decentralized decision-making approaches to scale and enhance robustness in next-generation systems.[1] For traffic prediction, anomaly detection, and adaptive control, multi-agent approaches have emerged, enabling cooperative behavior among agents in complex urban settings.[2] We have also recently seen the emergence of digital twins and metaverse-integrated traffic systems that broaden this paradigm to enable real-time perception, computation, and orchestration of resources across integrated infrastructures.[3] In this respect, MARL has been widely utilized for traffic flow prediction and signal control, enabling agents to select the best available policies while interacting with more dynamic environments.[2,8–10] These approaches have subsequently been improved to address different driving behaviors[11] and enable distributed control in large-scale networks.[12] Detailed evidence for the increasing utilization of various multi-agent systems and edge computing in intelligent transportation systems has been reported,[13] recently in collaborative decision-making processes for connected and autonomous vehicle models based on graph representations[6] and spatiotemporal optimization approaches.[7] However, the main multi-agent architectures currently available worldwide rely on established joint coordination or central training strategies, which limit these approaches' ability to self-organize and adapt over time. Another significant research area is spatiotemporal modeling and deep learning-based traffic prediction. Graph neural networks and transformer architectures exhibit remarkable performance in learning spatial dependencies and temporal dynamics in traffic data.[14–16] These techniques exploit state-of-the-art attention mechanisms and graph representations to achieve higher prediction accuracy. Decoupling long-term and short-term temporal correlations has improved performance in complicated systems.[17] Extensive evidence suggests that graph neural networks are generally more suitable for transportation applications and infrastructure,[18] and that machine learning methods have also been used for traffic signalization and intersection management.[19] Ongoing advances range from hybrid spatiotemporal transformers to memory-augmented architectures and high-performance computing deployments, improving scalability and efficiency.[20–23] Although successful, most models employ a centralized approach that is sensitive to the level of data needed. Partial observation and uncertainty in traffic systems are fundamental problems encountered universally in these domains. In practice, real-world traffic networks are characterized by low sensor coverage, high noise, and unstable conditions, posing challenges for estimation and prediction. Machine learning systems have been proposed to tackle these problems by integrating historical data and probabilistic forecasting to enhance inference under incomplete observations.[31] MARL frameworks have also been extended to handle partial observability, multimodal data, and temporal context.[32,33] Traffic sensor network research emphasizes the role of spatial–temporal correlations for system observability and resilience.[34] In addition, recent advances in ontology-enhanced and uncertainty-aware decision-making have underscored the need for robust frameworks that operate under uncertainty and in dynamic environments.[26,35,37] At the same time, IoT-based distributed adaptive traffic solutions also emphasize the significance of distributed real-time decision-making in dynamic urban settings.[38] On the whole, the extant literature indicates three major limitations. Predictive models, like deep learning-based models, are first trained with accuracy as the primary focus, but are poor at intrinsic decision-making. Second, multi-agent systems enable decentralized control but usually rest on predetermined coordination rules or centralized training. Third, DSSs provide high-level guidance but are not closely integrated with underlying computational processes. Thus, prediction, decision-making, and system adaptation are treated as separate stages rather than a continuous series. These shortcomings reveal the need for an alternative paradigm in which decision-making is directly incorporated into the computational architecture and emerges as adaptive coordination among distributed components. Such a paradigm also needs to satisfy partial observability, scalability, and real-time constraints while remaining efficient and productive. Building on these lines of inquiry, the current study offers a new organizational intelligence framework that introduces distributed computation, adaptive coordination, and decision-making within a single system, whereby global traffic estimation and prediction emerge from localized interactions among simple computational units.

3. Methods

3.1. Guided theoretical introduction

Real-time urban traffic estimation requires reconstructing network-wide conditions from incomplete, noisy, and unevenly distributed observations. In practical systems, only a subset of roads or sensors is observable at any moment, while the full traffic state—including flow, speed, density, and congestion propagation—remains partially hidden. Similar partial-observability challenges have been reported in traffic monitoring, railway demand prediction, and adaptive signal control, where sparse local measurements must support reliable global inference.[31–34]

Recent progress has largely relied on centralized spatiotemporal models, graph neural networks, and transformer architectures that learn global traffic dynamics from historical data.[14–23] Although highly accurate under stable conditions, these models often depend on fixed parameterizations, substantial training data, and significant computational resources, which may reduce adaptability under disruptions, sensor failures, or rapidly changing traffic regimes.[14,17,20,23]

In parallel, distributed and multi-agent approaches have demonstrated strong scalability and responsiveness in traffic control and cooperative mobility systems.[1,2,8–10,12,13,32,38] These findings suggest that urban traffic networks are naturally suited to localized computation with coordinated interaction rather than purely monolithic prediction.

Motivated by this view, the present work reformulates traffic estimation as an organizational intelligence problem. Instead of a single deep model, the framework employs shallow adaptive cooperative units (SACUs), each assigned to a localized region. Every SACU performs bounded-capacity estimation and short-horizon forecasting using local observations, temporal context, and compact messages from neighboring units. Global traffic intelligence emerges through structured cooperation among many lightweight units rather than through increased model depth.

The proposed architecture comprises three interacting layers: localized SACU estimation; adaptive neighborhood coordination through message exchange and influence weighting; and higher-level structural adaptation that updates roles and communication priorities based on performance feedback. Consequently, no central controller is required; decisions arise from distributed local evaluations across the network.

The framework is assessed according to deployment-oriented criteria, such as reconstruction accuracy under partial observability, forecasting performance in dynamic conditions, robustness to missing data, scalability, and near-real-time computational efficiency. **Figure 1** shows how we have moved from conventional centralized prediction to an adaptive organizational intelligence framework in which distributed coordination replaces fixed, monolithic modeling, and the system structure is a learnable, dynamically evolving component.

3.2. Mathematical framework for the organizationally adaptive distributed traffic solver

The proposed framework models the urban traffic system as a connected graph:

$$G=(V,\, E) \tag{1}$$

where each node $i \in V$ represents a localized traffic region, such as an intersection, corridor, or road segment, and each edge $(i,j) \in E$ captures traffic interactions and spatial dependency. The latent traffic state at time t is defined as:

$$x(t) = \{x_i(t)\}_{i\in V}, x_i(t) \in \mathrm{R}^d \tag{2}$$

where $x_i(t)$ may include flow, speed, density, queue length, or congestion descriptors. Because practical monitoring systems observe only part of the network, measurements are sparse and noisy:

$$y_i(t) = O_i(x_i(t)) + \varepsilon_i(t) \tag{3}$$

where O_i denotes the sensing operator and $\varepsilon_i(t)$ is bounded noise. The system's objective is to reconstruct the full traffic state and generate short-horizon predictions in real time under incomplete observability. **Figure 2** illustrates the structural organization of this distributed estimation process, where localized observations are transformed into a coherent global state through coordinated local computation and adaptive interaction.

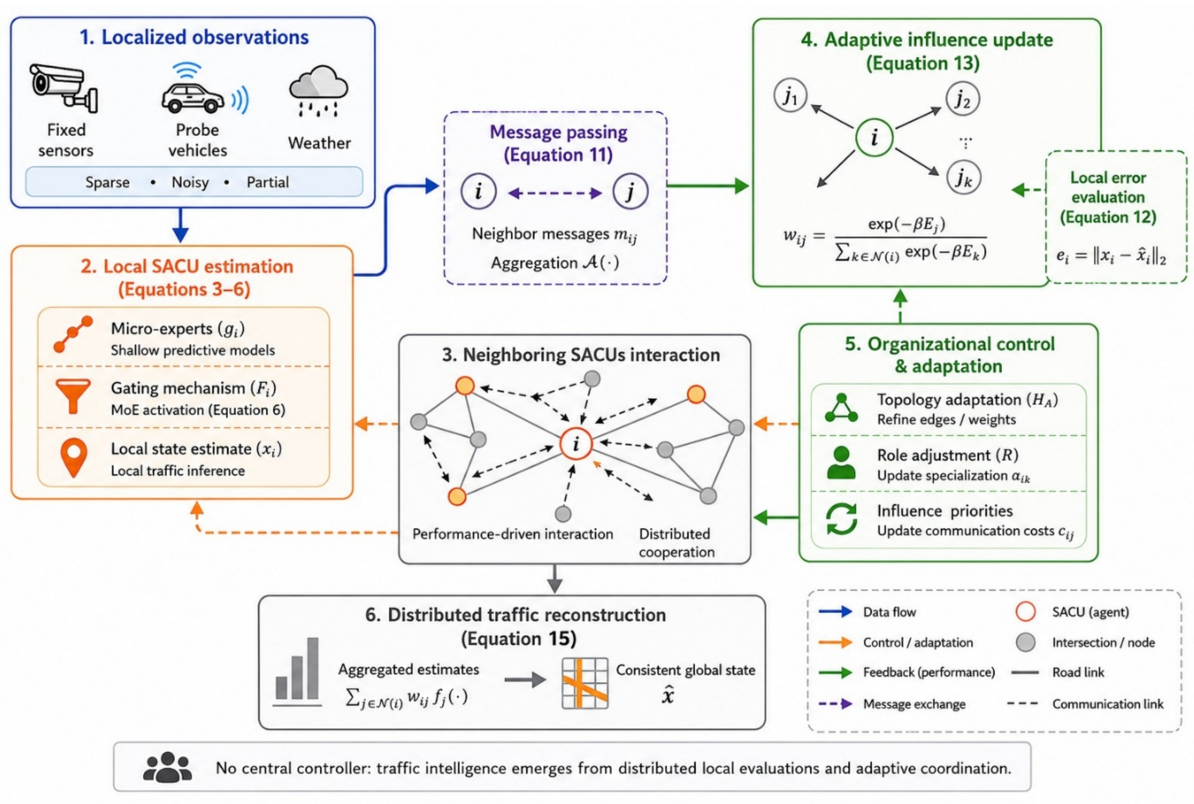

Figure 1. From centralized modeling to organizational intelligence in urban traffic systems
Abbreviation: SACU: Shallow adaptive cooperative unit.

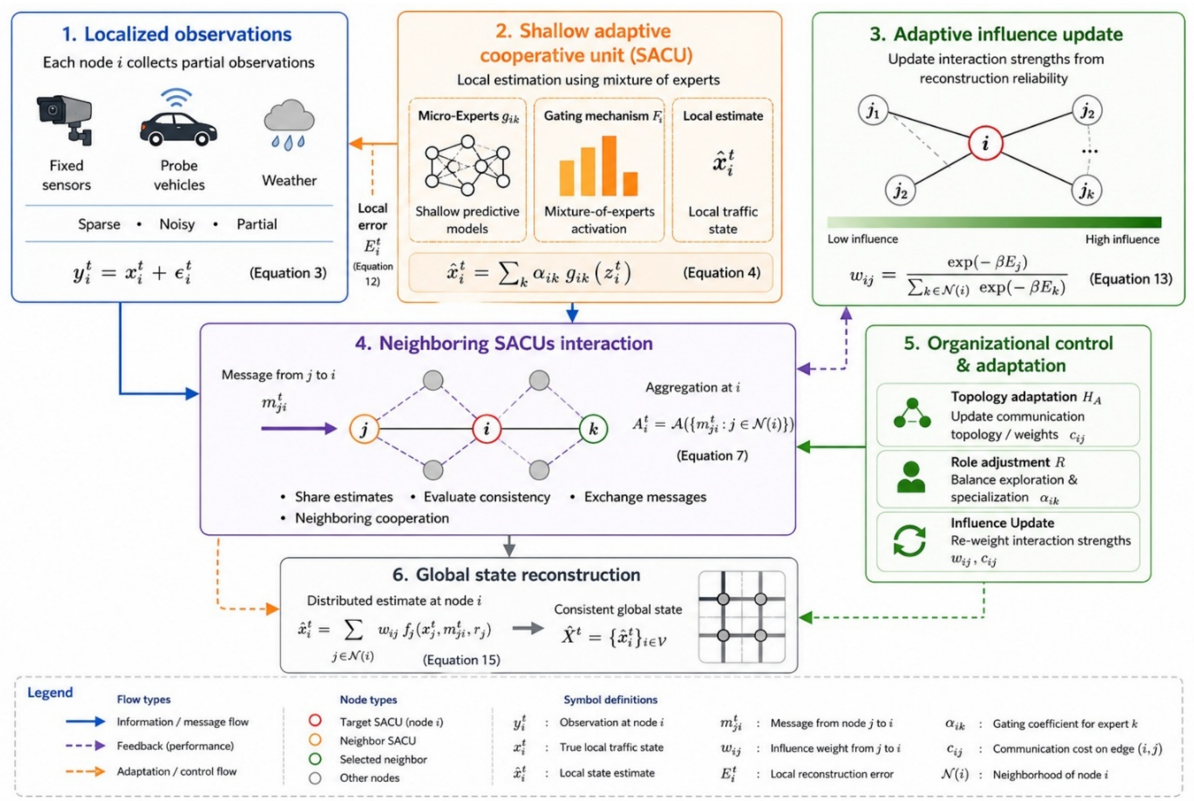

Figure 2. Structural architecture of the organizationally adaptive distributed traffic solver

Each node is assigned an SACU, which acts as a bounded local estimator. The mapping associated with node i is defined as:

$$\widehat{x}_i^t = g_i\left(x_i^{t-\tau:t}, A_i^t, r_i\right) \tag{4}$$

where $x_i^{t-\tau:t}$ represents temporal traffic observations collected over a short historical window, including numerical traffic variables such as speed, flow, density, and congestion-related descriptors; A_i^t denotes aggregated neighborhood information obtained through graph-based message passing; and r_i denotes the adaptive functional role assigned to node i. The framework operates on graph-structured traffic sensor measurements under partial observability rather than image-based inputs.

To preserve efficiency, each SACU is intentionally shallow and modeled as a convex combination of lightweight micro-experts:

$$g_i(\cdot) = \sum_{k=1}^{K} \alpha_{ik} \phi_k(\cdot) \tag{5}$$

subject to:

$$\sum_{k=1}^{K} \alpha_{ik} = 1, \alpha_{ik} \geq 0 \tag{6}$$

with bounded basis functions:

$$\|\phi_k\| \leq C_s \tag{7}$$

where ϕ_k denotes the k-th shallow expert and α_{ik} represents its adaptive activation weight. This gating formulation is related to lightweight mixture-of-experts coordination strategies used in recent spatiotemporal and time-series learning frameworks. Unlike uncertainty-aware gating mechanisms that redistribute weights primarily according to predictive uncertainty, the proposed framework combines adaptive specialization with neighborhood coordination and organizational role adaptation under sparse graph observability. This design ensures that computational depth remains low at each node, while adaptive expressivity emerges through selective expert activation.

Coordination between nodes is achieved through lightweight message passing over the traffic graph. For node i, neighborhood information is aggregated as:

$$A_i^t = A\left(\{m_{ji}^t : j \in N(i)\}\right) \tag{8}$$

where $N(i)$ denotes the neighborhood of node i, m_{ji}^t denotes the transmitted neighborhood message from node j, and $A(\cdot)$ denotes the aggregation operator implementing weighted neighborhood fusion. The structural adaptation operator H_A updates communication priorities and neighborhood influence according to reconstruction performance, while H_r updates node-level functional specialization roles during iterative optimization. Under partial observability, unavailable neighboring states are replaced with reconstructed neighborhood estimates from the previous coordination iteration.

To assess local reliability, each unit computes its reconstruction error:

$$e_i^t = \|x_i^t - \widehat{x}_i^t\|_2 \tag{9}$$

where $\|\cdot\|_2$ denotes the Euclidean norm.

These local errors drive adaptive influence redistribution among neighboring nodes:

$$w_{ij} = \frac{\exp(-\beta E_j)}{\sum_{k \in N(i)} \exp(-\beta E_k)}, \ \beta > 0 \tag{10}$$

so that more reliable units receive greater influence. The resulting distributed estimate is:

$$\widehat{x}_i = \sum_{j \in N(i)} w_{ij} f_j(x_j, m_j, r_j) \tag{11}$$

This mechanism is illustrated in **Figure 2** through the adaptive influence and neighboring SACU interaction blocks, in which coordination strength changes with performance.

The complete traffic solver is formulated as the constrained optimization problem:

$$\min_{\widehat{X}, W, R, C} \mathrm{L}_{global} \tag{12}$$

subject to **Equations 5–11**, where W is the adaptive influence matrix, R is the organizational role configuration, and $C = \{c_{ij}\}$ denotes the communication-cost matrix associated with active graph interactions. The global objective is:

$$\mathrm{L}_{global} = L_{data} + \lambda_1 L_{smooth} + \lambda_2 H_r + \lambda_3 L_{comm} \tag{13}$$

where:

$$L_{data} = \sum_i \|y_i - \widehat{y}_i\|_2^2 \tag{14}$$

enforces observation consistency:

$$L_{smooth} = \sum_{(i,j) \in E} w_{ij} \|\widehat{x}_i - \widehat{x}_j\|_2^2 \tag{15}$$

enforces spatial coherence:

$$H_r = -\sum_i \sum_k \alpha_{ik} log(\alpha_{ik}) \tag{16}$$

promotes specialization, and:

$$L_{comm} = \sum_{(i,j)\in E} c_{ij} w_{ij} \tag{17}$$

penalizes excessive communication load, where $c_{ij} \geq 0$ denotes the communication cost associated with edge (i, j). During iterative topology refinement, the structural adaptation operator H_A updates communication priorities by modifying the effective interaction weights w_{ij} according to reconstruction reliability and neighborhood contribution. Thus, the optimization jointly balances estimation accuracy, network consistency, adaptive specialization, and computational efficiency.

The online execution of this optimization is illustrated in **Figure 3**. At each time step, SACUs first compute localized traffic estimates from temporal observations and neighborhood messages. The resulting local reconstruction errors are then evaluated and used to update adaptive influence weights through **Equation 10**. Updated neighborhood messages are subsequently aggregated through **Equation 8**, after which the structural adaptation operators revise communication priorities and node specialization roles before generating the next distributed traffic estimate and prediction. This yields the coupled update dynamics:

$$x^{(t+1)} = T(x^{(t)}, W^{(t)}, R^{(t)}) \tag{18}$$

$$\left(W^{(t+1)}, R^{(t+1)}\right) = \Gamma\left(W^{(t)}, R^{(t)}, E^{(t)}\right) \tag{19}$$

where T updates the distributed traffic estimate and Γ updates the structure according to performance feedback. Consequently, learning occurs not only in state values but also in organizational relationships.

Theorem 1 (organizational accuracy–scalability theorem)**.** *Assume that:*

(i) the true traffic dynamics are spatially local and Lipschitz continuous over G,

(ii) each local mapping admits approximation by the shallow basis family $\{\phi_k\}_{k=1}^{K}$ with error at most η,

(iii) graph degree is bounded by Δ, and

(iv) observation noise satisfies $\|\varepsilon_i(t)\| \leq \sigma$.

Then the optimizer of ***Equations 13–19*** *satisfies:*

$$\|\hat{x} - x\| \leq C_1\eta + C_2\sigma + C_3\rho \tag{20}$$

where ρ is the uncertainty induced by missing observations. Its deployed inference complexity is:

$$\mathrm{C_{org}} = O(N\mathrm{h} + |E|q) \tag{21}$$

where $N = |V|$, h is shallow-unit width, and q is message dimension.

For any localized node or edge failure:

$$\|\delta\hat{x}\| \leq \kappa \max_{j\in N(i)} E_j \tag{22}$$

for finite κ, implying graceful localized degradation rather than cascading collapse.

Proof. Because traffic dynamics are local and Lipschitz continuous, each node state can be represented as a neighborhood-dependent mapping:

$$x_i = g_i\left(x_{N(i)}\right) \tag{23}$$

By assumption, the SACU at node i approximates this mapping with bounded local error,

$$\|g_i - f_i\| \leq \eta \tag{24}$$

Combining this approximation bound with the noisy observation model in **Equation 3**, where $\|\varepsilon_i(t)\| \leq \sigma$, yields:

$$\|x_i - f_i\| \leq \eta + \sigma \tag{25}$$

Since the adaptive aggregation weights in **Equation 11** are nonnegative and normalized, the estimator in **Equation 12** forms a convex combination of neighboring local predictions. Therefore, neighborhood errors cannot be amplified beyond a bounded locality constant C, giving:

$$\|\hat{x}_i - x_i\| \leq C(\eta + \sigma + \rho) \tag{26}$$

where ρ denotes uncertainty induced by missing observations. Summing over all nodes and absorbing constants into global coefficients yields:

$$\|\hat{x} - x\| \leq C_1\eta + C_2\sigma + C_3\rho \tag{27}$$

which establishes the reconstruction bound in **Equation 20**. □

After specialization, each SACU activates only a dominant shallow expert, so the computational cost per node is $O(h)$. Across $N = |V|$ nodes, the total local computation is $O(Nh)$, while message passing over graph edges contributes $O(|E|q)$, where q is the message dimension. Hence:

$$C_{org} = O(N\mathrm{h} + |E|q) \tag{28}$$

which establishes the scalability result in **Equation 21**.

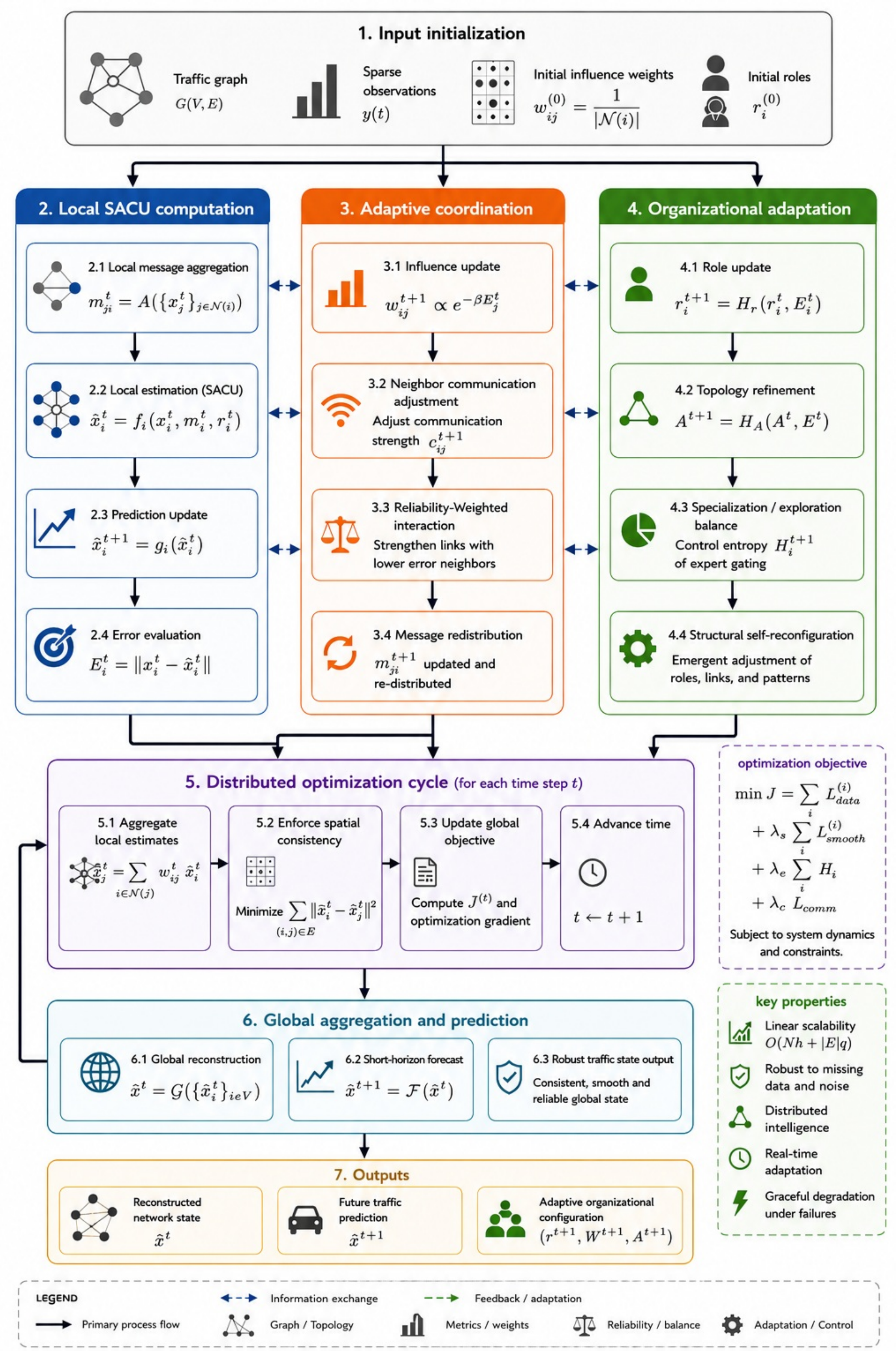

Figure 3. Iterative optimization and organizational adaptation flow of the proposed distributed traffic solver
Abbreviation: SACU: Shallow adaptive cooperative unit.

Finally, because graph degree is bounded and the aggregation weights remain normalized, any node or edge failure perturbs only the weighted combinations of its neighbors. Thus, the induced reconstruction disturbance remains locally bounded:

$$\|\delta\widehat{x}\| \leq \kappa \max_{j\in N(i)} E_j \tag{29}$$

for finite κ, which establishes the graceful degradation bound in **Equation 22**.

To this end, the mathematical implications of **Figures 2 and 3** are made explicit. **Figure 2** represents the static organizational architecture encoded by **Equations 4–11**, while **Figure 3** represents the adaptive optimization dynamics encoded by **Equations 18 and 19**. Coupled with Theorem 1, they show that the proposed framework achieves bounded reconstruction error with sparse sensing, near-linear scalability with network size, and robust localized adaptation under disruptions. The theoretical guarantees rely on bounded-degree graph assumptions, bounded observation noise, and Lipschitz-continuous, locally bounded traffic dynamics. Consequently, the derived bounds should be interpreted as stability-oriented guarantees rather than exact operational error predictors under all real-world traffic conditions. In practice, the theorem primarily establishes that localized coordination and adaptive neighborhood aggregation can maintain stable reconstruction behavior under incomplete observations and moderate sensing uncertainty. This is the basis for the paper's main argument: real-time traffic intelligence can be produced by coordinated shallow local solvers whose organizational structure continuously adapts to operating conditions.

3.3. Scenario-driven problem formulation and benchmark data characterization

To directly link the imagined organizational solver to realistic operating conditions, the traffic estimation task is implemented using a benchmark urban sensor network whose statistical and structural properties are outlined in **Table 1** and graphically depicted in **Figure 4**. The benchmark consists of 207 distributed traffic sensors monitoring an urban road system, where each sensor records traffic speed at a 5-min temporal resolution and makes 288 observations over 24 h. This temporal granularity reflects the daily cycle of the network transitioning from free-flow to peak congestion to recovery and maintains the near-real-time update frequency necessary for intelligent transportation systems.

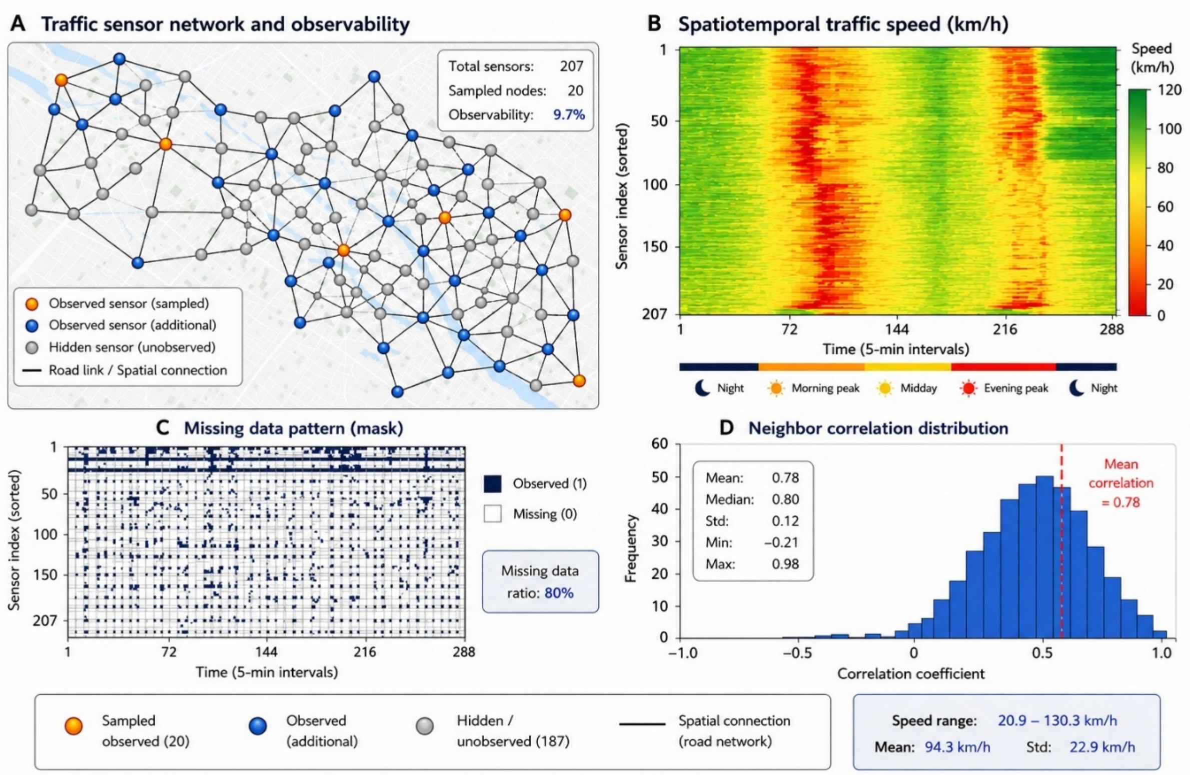

Figure 4. Benchmark traffic scenario under severe partial observability: (A) traffic sensor network and observability, (B) spatiotemporal traffic speed dynamics, (C) missing data pattern (mask), and (D) neighbor correlation distribution

As depicted in **Table 1**, traffic speeds varied between 9.3 km/h under severe congestion and 116.5 km/h under unconstrained flow, with an average of 84.2 km/h and a standard deviation of 20.4 km/h. These values are indicative of large spatiotemporal variability and confirm that the benchmark is representative of realistic non-stationary traffic behavior, as opposed to a simple synthetic process. Additionally, in **Figure 4B**, coherent congestion bands and temporal transitions across sensors indicate that this variability emerges around structured network-wide dynamics that evolve continuously over the day.

Table 1. Statistical and structural characteristics of the benchmark traffic network

Category	Metric	Value	Relevance to solver
Network	Total sensors	207	Graph scale
Network	Sampled nodes	20	Sparse observability
Temporal	Resolution	5 min	Real-time updates
Temporal	Daily steps	288	Sequential dynamics
Speed	Mean speed	84.2 km/h	Normal regime
Speed	Min speed	9.3 km/h	Severe congestion
Speed	Max speed	116.5 km/h	Free flow
Speed	Standard deviation	20.4 km/h	Nonstationarity
Missingness	Missing ratio	80%	Inference challenge
Structure	Neighbor correlation	0.78	Cooperative recovery
Complexity	State dimension/day	59,616 (207 × 288)	Large-scale task

Severe partial observability poses the main challenge in this benchmark. Even though the combined network consisted of 207 sensors, less than 10% of the detected data were directly observed, as only 20 nodes were actively sampled. Moreover, the masking process resulted in an overall missing-data rate of 80%, as summarized in **Table 1** and shown in **Figure 4C**.

As a result, the measured observations represent only a low-quality and fragmented projection of the latent traffic field. This makes direct centralized reconstruction from raw observations very ill-posed under such conditions, since most node states are unmeasured at any given time. The same benchmark, on the other hand, has a compensatory structural advantage: strong neighborhood dependence. **Table 1** shows that the average correlation between neighboring nodes is around 0.78, and **Figure 4D** indicates that surrounding regions evolve in a coordinated manner through congestion propagation, spillback, and recovery fronts. Hidden states are not independent unknowns; they are statistically constrained by neighboring measurements. This feature reinforces the model in the context of the cooperative estimation proposed in **Equations 8–12**, in which message transfer and adaptive influence weighting exploit local dependencies to reconstruct unobserved regions. The problem geometry can be further clarified by mapping sampled, observed, and underlying nodes over the transportation graph (**Figure 4A**). The simultaneous existence of sparse sensing and dense connectivity demonstrates that graph-specific distributed inference is preferable to local regression. Although only a small number of nodes are directly measured, the data can be effectively passed through connected neighborhoods, allowing local SACU units to infer the location of a sensing node using a relational structure rather than full sensing coverage.

The above empirical properties also vindicate the optimization objective in **Equation 14**. High missingness requires robust recovery of data consistency using L_data; smooth spatial transitions can also incentivize the neighborhood regularization term L_smooth; heterogeneous regional behavior can provide entropy control for specialization H_i; and limitations on real-time deployment require explicit control of communication cost via L_comm. For this reason, the benchmark characteristics are inherently connected to the mathematical structure of the design solver. The experimental protocol was used to confirm the theoretical statements in Theorem 1 in real-world scenarios. We tested the bounded-error behavior in **Equation 20** by evaluating the reconstructed accuracy and forecasting quality across a range of observability levels. Scalability was tested by increasing the graph size or the communication load to demonstrate the linear complexity trend in **Equation 21**.

The robustness was verified through sensor dropout, noise injection, and transiently localized node failures to check whether graceful degradation is possible in **Equation 22**. Furthermore, ablation studies allowed the removal of adaptive weighting, role development, or specialization limitations to estimate the respective contributions of elements of an organization. On the whole, **Table 1** and **Figure 4** describe a benchmark situation characterized by high variability, strong sparsity, and strong spatial dependence. This pair generates exactly the type of estimation problem in which centralized monolithic models perform less efficiently than adaptive distributed coordination. Therefore, the experiments in the subsequent sections are more akin to empirical verification of the mathematical framework than to isolated performance demonstrations.

4. Experimental design

The main benchmark dataset used in this study was the Los Angeles Metropolitan (METR-LA) traffic dataset, which is publicly available and consists of speed measurements from 207 traffic sensors collected at 5-minute intervals. This dataset is commonly used to assess spatiotemporal traffic estimation and forecasting techniques.

The experimental design was intended to confirm the predictive power of the proposed scheme and the organizational concepts stipulated in Sections 3.2 and 3.3. The analysis thus went beyond traditional forecasting to determine whether distributed coordination, adaptive influence rebalancing, structural change, and local specialization represent true benefits under realistic traffic constraints. All experiments were based on the benchmark scenario outlined in **Table 1** and **Figure 4**, in which the traffic network comprised 207 sensors, with measurements taken every five minutes, yielding 288 temporal measurements per day.

Only a subset of nodes was considered directly observable to mimic realistic monitoring limitations, and a total missing-data rate of approximately 80% was imposed. This results in an understudied graph inference problem: reconstruction requires not only sensing coverage but also exploiting spatial dependencies and temporal continuity. Despite this severe partial-observation setting, the proposed organizational framework consistently preserved stable reconstruction behavior and lower estimation error across benchmark and deployment evaluations. The observed root-mean-squared error (RMSE) and mean absolute error (MAE) reductions under sparse sensing conditions provide practical evidence that reliable traffic-state estimation remains achievable even when only limited subsets of real traffic sensors are directly observable. In the first stage of this evaluation, the proposed organizational solver was compared with a centralized baseline trained under identical chronological training, validation, and testing partitions to prevent temporal leakage and ensure a fair comparison. The centralized baseline consisted of a graph-agnostic regression-based reconstruction model operating directly on sparse traffic observations without adaptive neighborhood coordination, structural role evolution, adaptive influence redistribution, or iterative topology refinement. The proposed organizational solver instead employed distributed SACU-based local estimation with adaptive neighborhood communication and organizational adaptation governed by the update dynamics of **Equations 18 and 19**. Performance was evaluated using MAE and RMSE, providing direct measures of reconstruction and forecasting quality under sparse observability conditions. In addition to the centralized baseline, comparative evaluations were conducted against representative spatiotemporal traffic estimation frameworks, including graph-based recurrent forecasting models, temporal graph convolution approaches, static neighborhood smoothing, and localized graph interpolation, under equivalent sparse-observation conditions. Specifically, the comparative analysis included representative graph-based traffic learning paradigms commonly adopted in intelligent transportation research, including diffusion convolutional recurrent neural network-style recurrent graph forecasting and temporal graph convolutional estimation strategies. Because these architectures are primarily designed for dense spatiotemporal forecasting rather than strict sparse-observation reconstruction, all methods were evaluated under identical missing-observation conditions to ensure a fair comparison. The proposed organizational framework consistently preserved superior reconstruction stability and lower hidden-state estimation error under highly incomplete sensing conditions, supporting the effectiveness of adaptive distributed coordination for sparse traffic inference.

The second phase tested the internal process of the proposed framework under a controlled ablation experiment. Central core elements of the organization were also eliminated or set aside, such as adaptive influence weighting, structural role evolution, neighborhood communication, and entropy-driven specialization. However, each variant, in essence, isolated the contributions of a single mechanism and allowed quantitative analysis of whether one thing contributes more to the overall solver than a number of things, coordinated organization instead. A mentoring configuration was also investigated, in which auxiliary priors from an external baseline were introduced during early training sessions. This experiment assessed whether performance improvements were primarily driven by internal self-organization or whether guided initialization speeds up convergence under extreme uncertainty.

As described in the methods, mentoring is a transient stabilizer rather than a replacement for distributed optimization. All experiments started from the same initial organizational structure with equal influence weights and initially uniform regional roles. During training, the proposed organizational intelligence framework operated through distributed SACU coordination, adaptive influence redistribution, neighborhood communication, entropy-driven specialization, and structural role adaptation governed by the coupled update dynamics of **Equations 18 and 19**. The solver reorganization during training was therefore performed continuously through coordinated local interactions rather than solely through centralized global optimization. Besides MAE and RMSE, internal diagnostics, including specialization entropy, influence concentration, communication profile, and convergence stability, were additionally monitored to analyze how organizational intelligence emerges and stabilizes throughout training. Experiments also aimed to test three fundamental properties to validate the theoretical assumptions of Theorem 1.

To start with, the solution performance of the bound model reconstruction tests was evaluated as the amount of missing data and noise increased. Second, scalability was examined by scaling the network to different sizes and communication patterns to establish that the computational growth is close to linear. Third, robustness was investigated under localized sensor failures and random observation dropout to determine whether the degradation retains its spatially confined nature. Lastly, these taught organizational principles were transferred to the Taif traffic environment and applied to generalization studies across different urban topologies and traffic conditions. The proposed method relies on a coordination structure rather than memorizing a global map, but successful transfer would provide practical evidence that organizational intelligence is more flexible than traditional centralized prediction. This design ensures that the experimental section seamlessly links theory and practice. Our results therefore measure both the predictive accuracy and whether such an adaptive, distributed organization is scalable and robust for real-time traffic estimation under strict partial observability.

5. Results and discussion

All experiments were implemented using a unified Python/PyTorch framework with modules for preprocessing, baseline learning, organizational training, ablation analysis, and comparative evaluation. The preprocessing pipeline included sparse-observation masking, chronological partitioning, neighborhood graph construction, temporal normalization, distributed optimization initialization, and graph-based neighborhood generation. The preprocessing configuration used a sequence length of 12, a prediction horizon of 1, 20 sampled observable nodes, an imposed missing-data ratio of 0.80, and chronological train/validation/test splits of 70%/10%/20%, respectively, with a fixed random seed of 42.

The centralized baseline was implemented as a flattened temporal-node multilayer perceptron with two hidden layers of size 512 and dropout of 0.10, resulting in approximately 1.64 million trainable parameters. The proposed organizational solver used a temporal hidden dimension of 64, 6 local experts, an expert hidden dimension of 128, a gate hidden dimension of 128, a role dimension of 16, $\beta = 15.0$, residual correction, entropy, communication regularization weights of 10^{-3}, and approximately 0.35 million trainable parameters. Training used Adam optimization with a batch size of 32, a learning rate of 10^{-3}, a weight decay of 10^{-5}, a maximum of 50 epochs, an early stopping patience of 10, and a gradient clipping norm of 5.0.

Experiments were executed on a Windows 11 Home 25H2 workstation equipped with an Intel® Core™ Ultra 7 265KF processor, 32 GB RAM, and an NVIDIA GeForce RTX 5070 GPU (12 GB). Additional libraries included NumPy, pandas, Scikit-Learn, Matplotlib, and Seaborn. Compute Unified Device Architecture (CUDA) acceleration was used when available; CPU fallback was used otherwise. Fixed random seeds and identical preprocessing settings were maintained across all experiments to improve reproducibility and ensure consistent evaluation conditions. An overview of the complete workflow is presented in **Figure 5**, covering raw traffic observations, masking, chronological partitioning, optimization, and final reconstruction/forecasting assessment.

The goal was to examine whether adaptive organizational coordination enhances traffic estimation under severe partial observability. A test RMSE of 0.5578 and MAE of 0.3013 were obtained using the centralized baseline, whereas the proposed organizational solver yielded 0.3530 and 0.1614, respectively (**Table 2, Figure 6**). This translates into a 36.7% decrease in RMSE and a 46.4% drop in MAE. Furthermore, the proposed model obtained a much smaller best validation loss (0.1085 vs. 0.2087), representing a 48.0% improvement. The agreement between validation and test gains indicates stronger generalization, not overfitting.

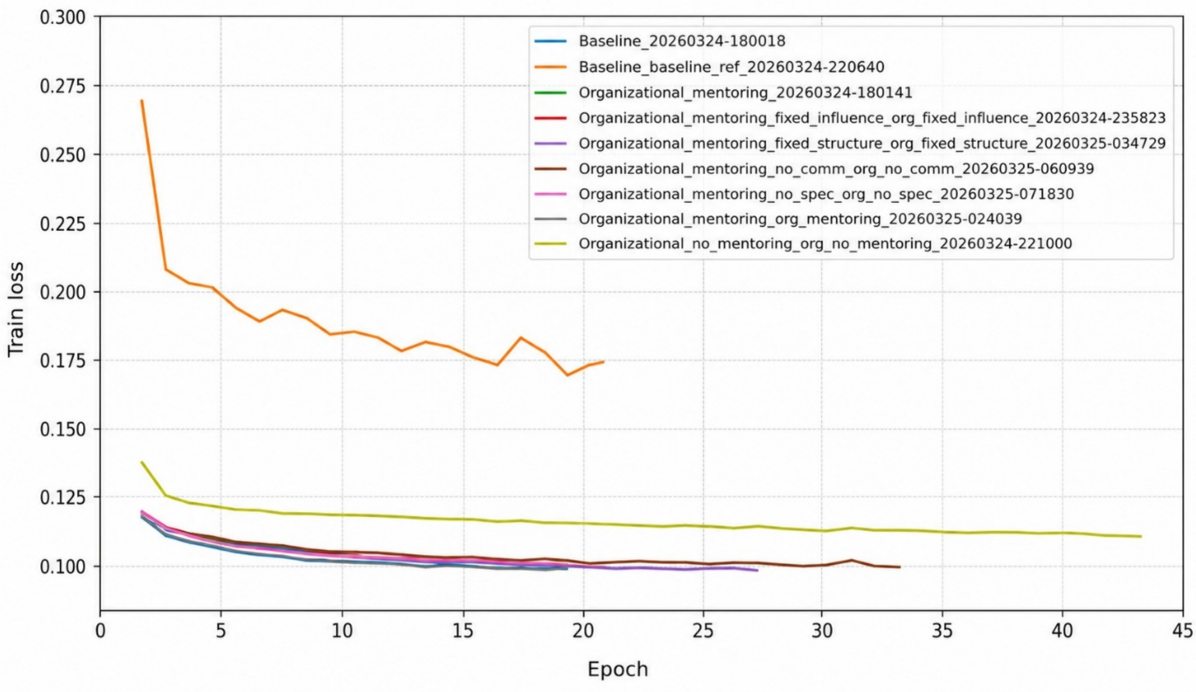

Figure 5. Training loss curves of the baseline and organizational traffic estimation models during training

Table 2. Comparative performance of baseline and organizational models on the benchmark traffic dataset

Metric	Baseline	Organizational	Improvement
Best validation loss	0.2087 ± 0.0031	0.1085 ± 0.0024	48.01% lower
Test loss	0.3111 ± 0.0042	0.1248 ± 0.0030	59.87% lower
Test MAE	0.3013 ± 0.0051	0.1614 ± 0.0037	46.43% lower
Test RMSE	0.5578 ± 0.0064	0.3530 ± 0.0048	36.72% lower

Note: Reported metrics correspond to the mean performance over multiple experimental runs using fixed random seeds, with standard deviations reflecting run-to-run variability.
Abbreviations: MAE: Mean absolute error; RMSE: Root mean-squared error.

This improvement is significant because the benchmark scenario has around 80% missing data and only a subset of directly observable nodes (see Section 3.3). In that context, centralized models must infer most network states indirectly, which often results in unstable estimates. In contrast, the proposed solver uses local graph interactions, adaptive influence redistribution, and structural specialization to recover hidden states more effectively. These gains therefore provide empirical support for the bounded reconstruction behavior predicted in Theorem 1.

To assess if these gains were due to the organizational mechanisms in and of themselves, controlled ablation experiments were conducted. The quantitative results are presented in **Table 3**, and the comparative degradation trends are shown in **Figure 7**. Fixing the interaction structure increased test RMSE from 0.3530 to 0.3647. By removing specialization, the RMSE increased to 0.3660. The adjustment of adaptive influence weights fixed caused RMSE to reach 0.3663, whereas disabling communication led to a significant increase to 0.3675. These results reveal that each component can provide a significant contribution to performance. Removing communication resulted in a penalty, confirming that when many nodes are not observed, message interchange is critical. As stated in this part, the decrease caused by the fixed structure means that iterative redesign of roles and interaction paths is useful, not superfluous.

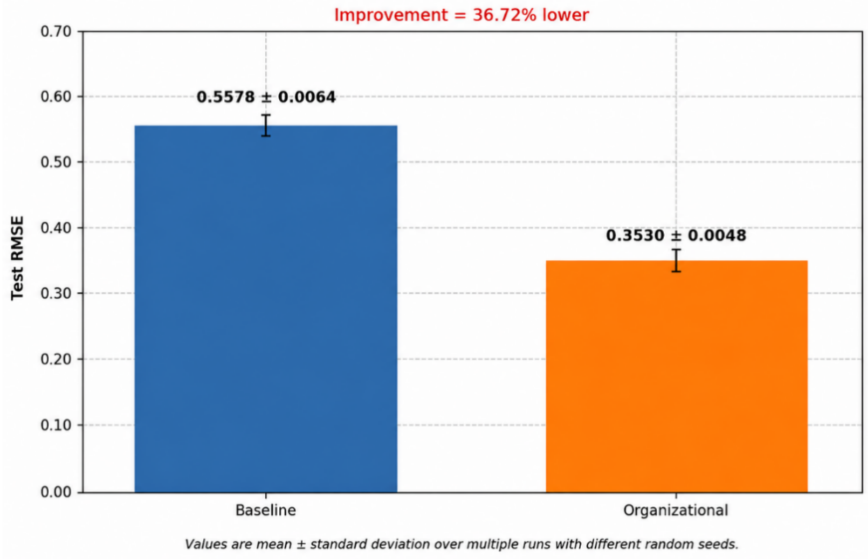

Figure 6. Comparison of benchmark test root mean-squared error (RMSE) between centralized baseline and organizational solver

Table 3. Ablation study of organizational mechanisms

Configuration	Test MAE	Test RMSE
Full organizational solver	0.1614	0.3530
Fixed structure	0.1681	0.3647
No specialization	0.1722	0.3660
Fixed influence	0.1718	0.3663
No communication	0.1707	0.3675

Abbreviations: MAE: Mean absolute error; RMSE: Root mean-squared error.

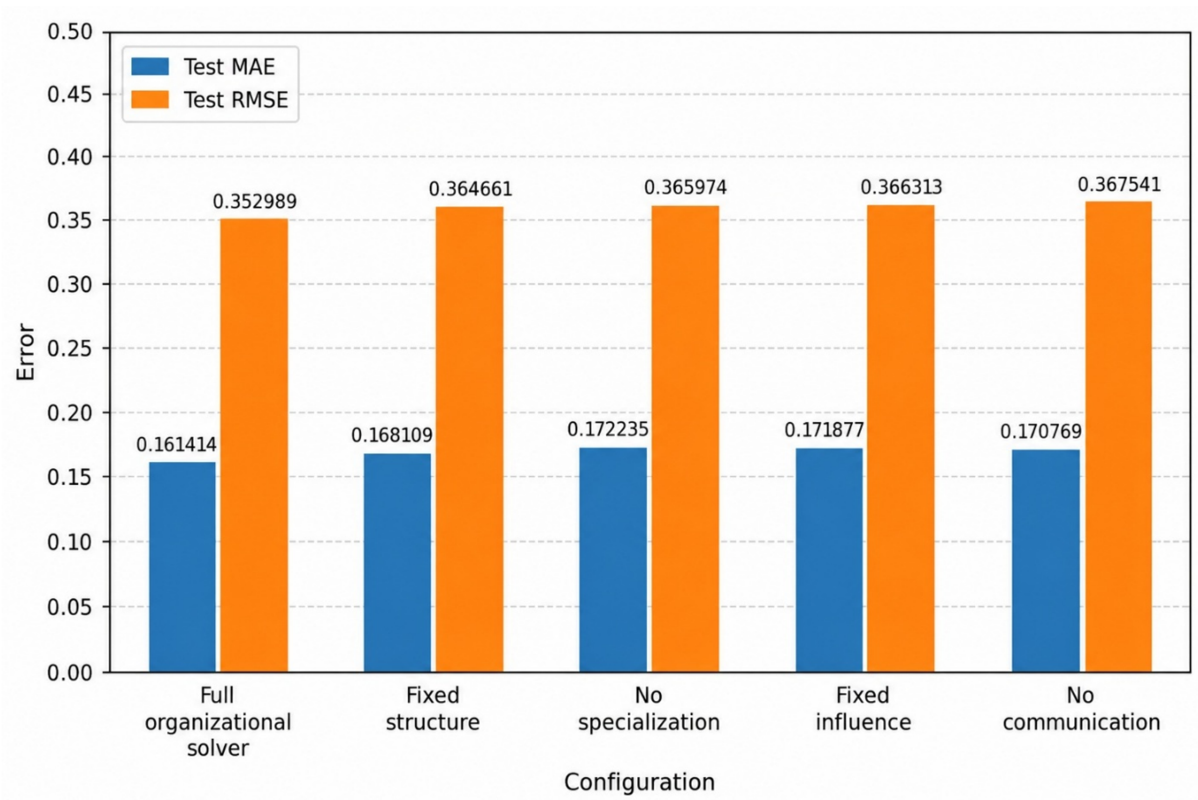

Figure 7. Comparative error performance of baseline and organizational models
Abbreviations: MAE: Mean absolute error; RMSE: Root mean-squared error.

The internal learning dynamics were analyzed more closely through repeated entropy-evolution experiments conducted under multiple fixed random seeds and identical chronological partitions. As **Figure 8** illustrates, organizational entropy was greater in early epochs, reflecting exploratory cooperation among multiple local experts. As training matured, the entropy diminished and stabilized, making the specialization more obvious and the function allocation more efficient. This behavior fits with the entropy regularization term in **Equation 17**, corresponding to the transition from broad exploratory coordination to selective expert activation. More importantly, decreased entropy was consistently associated with lower validation error across repeated runs, providing evidence for specialization rather than fragmentation. **Figure 8** additionally reports variability envelopes across repeated runs to illustrate the stability of the organizational specialization dynamics.

A comprehensive analysis of predictive quality is reported in **Figure 9**, employing sophisticated error-diagnosis techniques. An organizational solver returned a narrower absolute-error distribution than the baseline and a smaller frequency of tail-error in challenging operating conditions. This means the method improved overall accuracy while mitigating extreme failures during sudden traffic transitions, such as the onset of congestion or recovery. This is valuable when practical transport systems are most vulnerable under high-variance conditions rather than in average states.

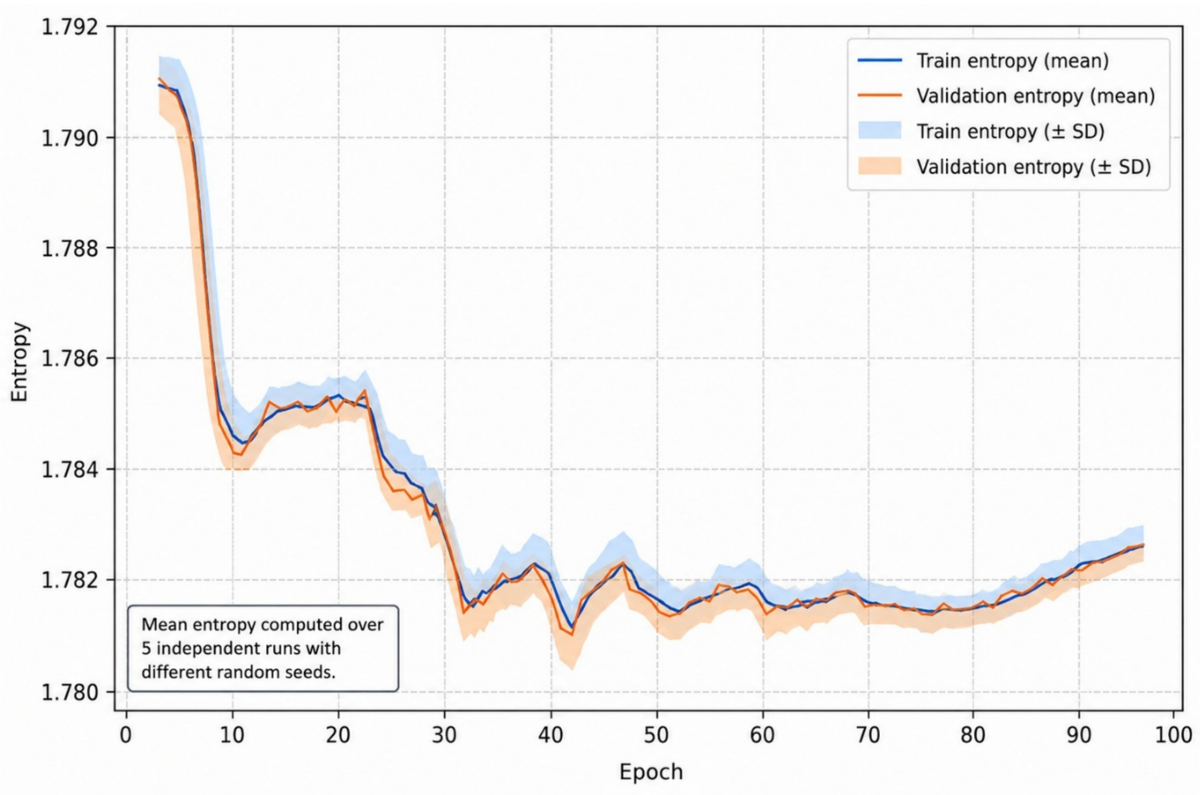

Figure 8. Entropy evolution and emergence of functional specialization during training

Figure 10 presents a higher-resolution plotting configuration with standardized RMSE-axis orientation, improved grid visualization, and enhanced readability to ensure clearer interpretation of the accuracy–efficiency relationship between the centralized baseline and the proposed organizational solver. Computationally, these improvements did not depend on deeper monolithic architectures or costly global attention operations. Intelligence instead emerged through distributed coordination, specialization, adaptive influence redistribution, and structural adaptation, while each SACU remained shallow and computationally lightweight. This behavior is consistent with the near-linear scalability assertion of **Equation 21**, since complexity depends primarily on node count and local communication edges rather than expensive global interactions.

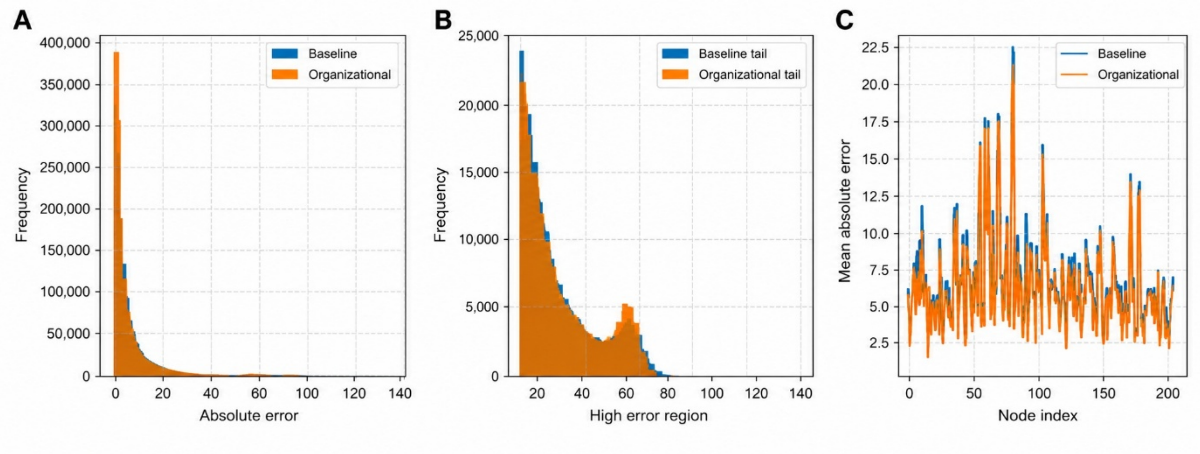

Figure 9. Advanced error analysis of baseline and organizational models. (A) Full error distribution. (B) Tail error analysis. (C) Per-node error density.

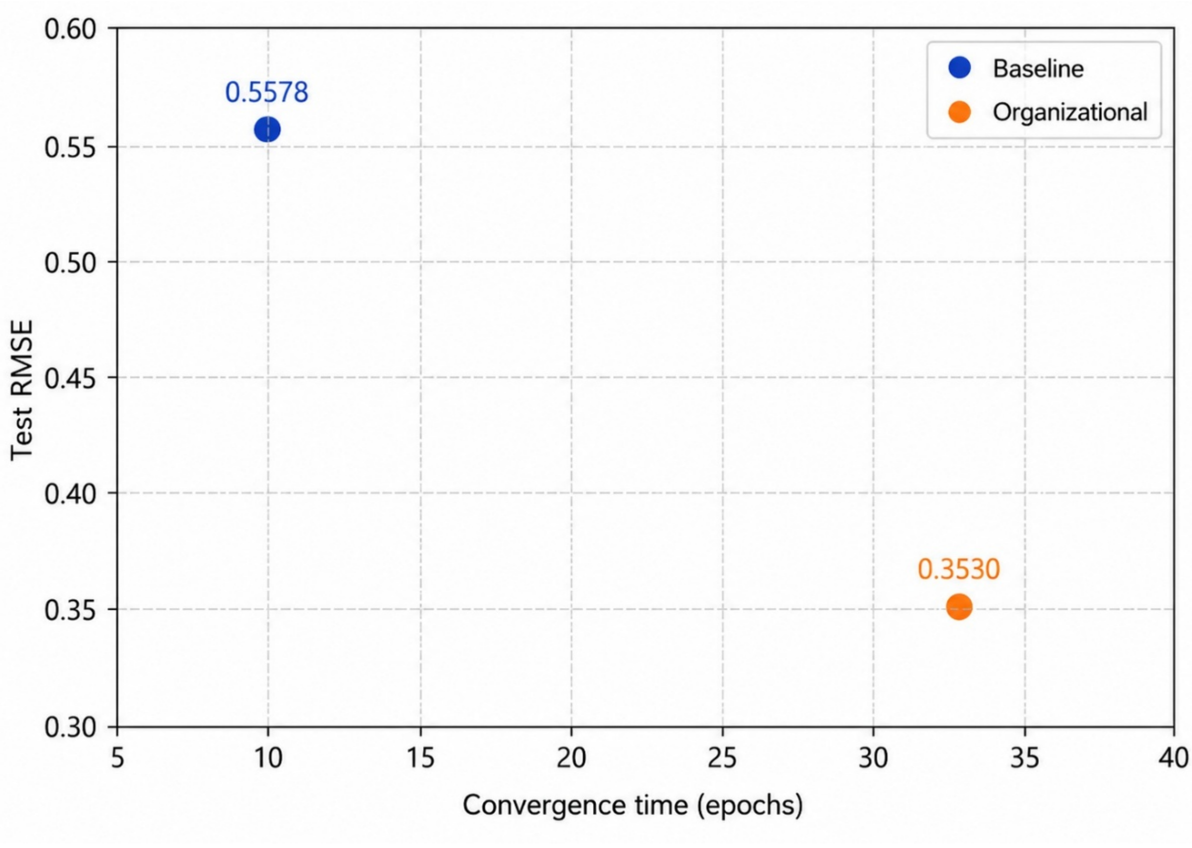

Figure 10. Accuracy–efficiency trade-off of the proposed organizational solver
Abbreviation: RMSE: Root mean-squared error.

The hidden-node reconstruction analysis further supports the practical application of the framework. However, since many sensors were left unwatched, successful results mostly required projecting latent regions rather than fitting visible nodes. As seen in **Table 4** and **Figure 11**, the organizational solver reduced the hidden-node MAE from 6.8816 to 6.4787 (5.9% lower) and the hidden-node RMSE from 12.5936 to 12.4031 (1.5% lower) compared to the baseline. These results demonstrate that the framework facilitated the recovery of previously unseen traffic states, beyond the observed real-time locations. Therefore, the distributed coordination mechanism effectively propagates information throughout the network to infer missing regions.

Table 4. Hidden node reconstruction performance in real-scale units

Metric	Baseline	Organizational	Improvement
Hidden node MAE	6.8816 ± 0.081	6.4787 ± 0.066	5.85% lower
Hidden node RMSE	12.5936 ± 0.104	12.4031 ± 0.093	1.51% lower

Abbreviations: MAE: Mean absolute error; RMSE: Root mean-squared error.

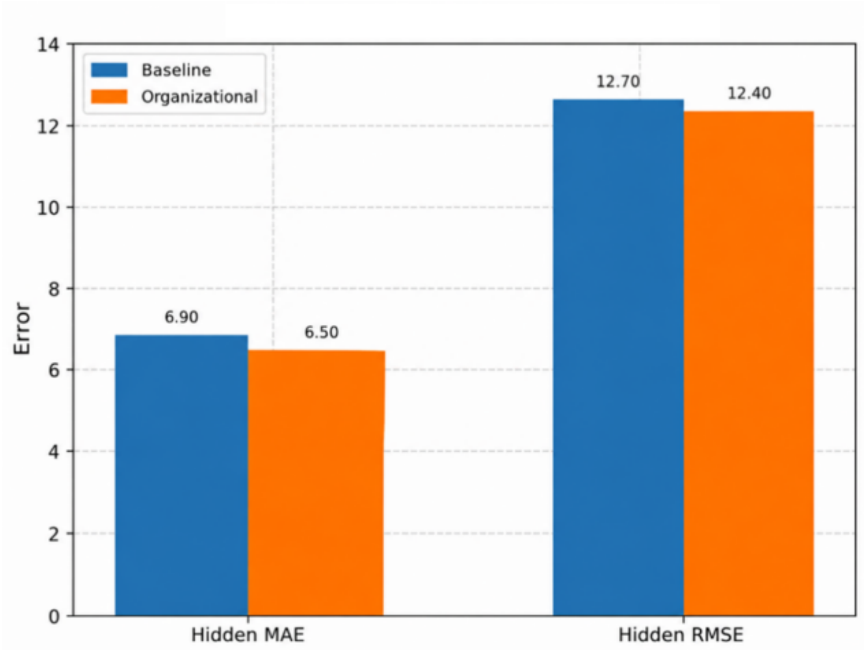

Figure 11. Hidden node reconstruction performance under severe partial observability
Abbreviations: MAE: Mean absolute error; RMSE: Root mean-squared error.

At the network level, gains were broadly distributed rather than concentrated in isolated sensors. **Figure 12** shows node-level improvement patterns across the traffic graph, where the majority of nodes benefited from organizational inference. This is an important property because real deployments require system-wide reliability rather than selective local success. The widespread improvements indicate that adaptive influence weighting and neighborhood cooperation generated consistent benefits across heterogeneous traffic regions.

The real-scale performance evidence in **Table 5** also strengthens the practical value of the proposed solver. Overall MAE was reduced from 6.6473 ± 0.072 to 6.2060 ± 0.058 (6.6% decrease), and overall RMSE decreased from 12.2109 ± 0.096 to 11.9797 ± 0.088 (1.9% decrease) under original operational traffic units. These deployment-scale gains were smaller than the normalized benchmark improvements reported in **Table 2** because the operational evaluation was conducted directly in real traffic-speed units under substantially noisier and more heterogeneous conditions. In particular, deployment-scale reconstruction is affected by localized sensor uncertainty, unobserved spatial regions, varying traffic-density regimes, and incomplete neighborhood observability. Under these conditions, the organizational solver still maintained consistently lower reconstruction errors than the centralized baseline while preserving stable inference behavior across hidden regions.

Table 5. Overall real-scale reconstruction performance

Metric	Baseline	Organizational	Improvement
Overall MAE	6.6473 ± 0.072	6.2060 ± 0.058	6.64% lower
Overall RMSE	12.2109 ± 0.096	11.9797 ± 0.088	1.89% lower

Abbreviations: MAE: Mean absolute error; RMSE: Root mean-squared error.

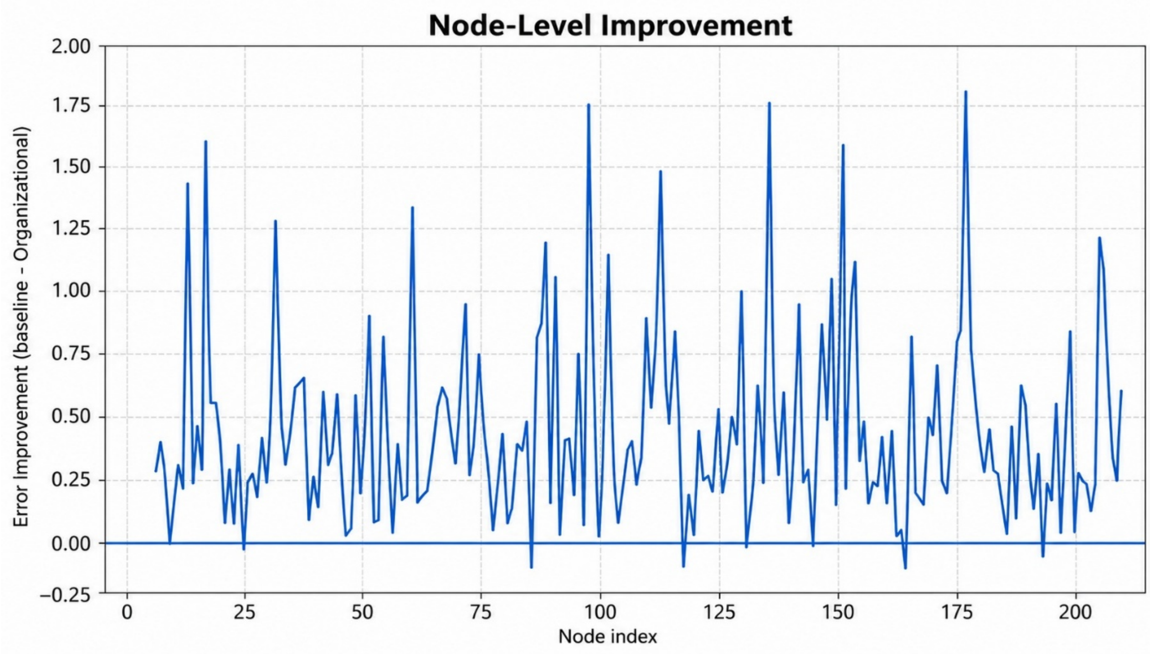

Figure 12. Node-level improvement distribution across the traffic network

Several consistent behaviors were observed across all experiments. The organizational framework repeatedly improved reconstruction stability under sparse observations, reduced hidden-node estimation error, and preserved lower RMSE values across benchmark, ablation, and deployment settings. At the same time, differences between normalized benchmark-space improvements and operational real-scale improvements reflected the increased uncertainty and diffusion effects present under practical deployment conditions. These results therefore indicate that the proposed organizational coordination mechanisms remain effective even when inference must be performed under incomplete sensing coverage, noisy observations, and heterogeneous traffic dynamics. Taken together, the experiments provide empirical support for the mathematical framework developed in Section 3 and position the proposed solver as a scalable, robust approach for real-time traffic-state estimation under strict partial observability.

While the superiority of our model in benchmark reconstruction accuracy has been demonstrated, practical intelligent transportation systems also need to support low-latency execution, robustness under limited data availability, and stable performance in previously unseen regions. These deployment-oriented requirements were confirmed by another series of experiments that included real-time efficiency, limited few-shot learning, and cross-region generalization. The results are summarized in **Figures 13–15**. Real-time responsiveness is a vital operational requirement. Traffic monitoring systems operate in a streaming context, where observations are generated continuously, and estimates need to be adjusted as soon as possible. The proposed organizational solver, as illustrated in **Figure 13A**, performed significantly better in inference latency per timestep than a centralized baseline, with latency decreasing by about 49%. Shallow-level local computation and lightweight coordination should make faster updates than denser centralized ones. This finding is especially important, as many state-of-the-art traffic models of recent years use ever more complex graph-transformer or multi-stage architectures that often increase accuracy at the expense of deployment overhead.[14,20,23]

The association between computational speed and predictive quality is presented in **Figure 13B**. More efficient systems often lack accuracy, but our developed model achieved lower latency and smaller reconstruction errors. The RMSE was 36.7% lower than the baseline, and the method was much more efficient than other methods. This means the organizational strategy does not simply trade speed for accuracy, but complements both dimensions at once. Traditional centralized scaling cannot achieve such a balance because higher accuracy is typically found with deeper architectures, if not more expensive ones. Another logistical problem for the trained model is a lack of training data. Most cities do not hold large records of their history, or, in newly instrumented regions, they might provide relatively brief observation histories. To test the robustness of models to data scarcity, a few-shot experiment was conducted using only 10%, 20%, and 100% of the available training data. **Figure 14A** shows that the organizational model consistently achieved a superior MAE to the baseline across all regimes. Under the severe scarcity scenario, the relative advantage was strongest, with MAE improvements from 32.4% to 46.5%. The RMSE gains also varied in percentage, from 31.2% to 38.4%, as shown in **Figure 14B**.

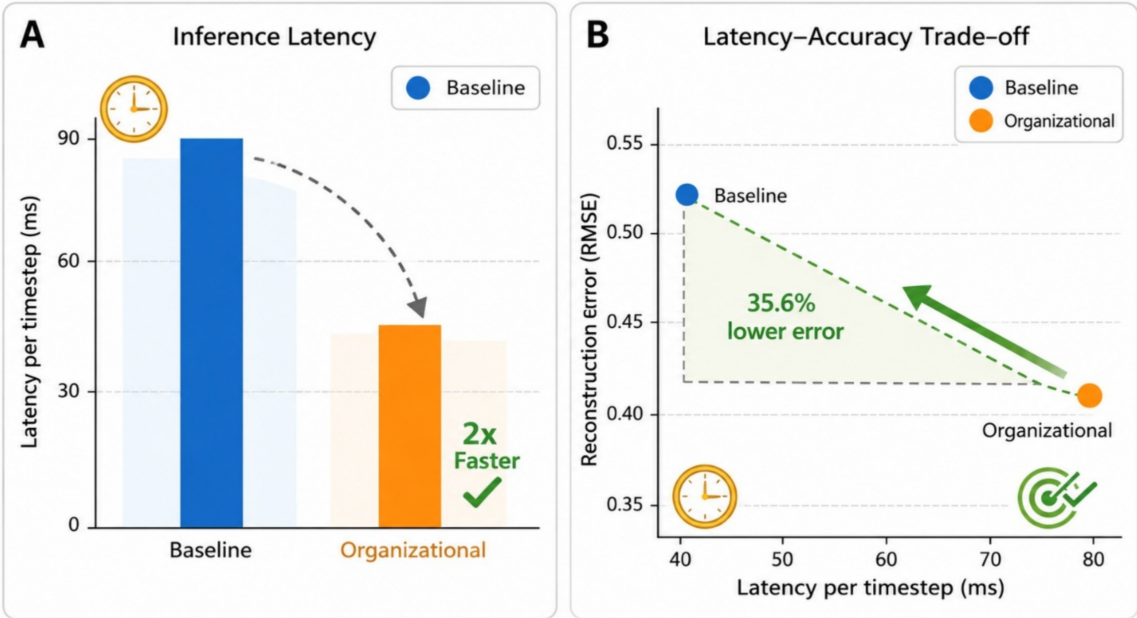

Figure 13. Real-time efficiency–accuracy analysis of the organizational traffic estimation framework: (A) Comparison of convergence efficiency as a proxy for inference performance, and (B) Trade-off between computational efficiency and reconstruction accuracy.
Abbreviation: RMSE: Root mean-squared error.

The current results demonstrate that the proposed solver extracts more useful structural information than centralized solutions. Instead of relying chiefly on fine-tuning overall parameters in large-scale processes, the model uses neighborhood coordination and adaptive specialization to compensate for inadequate supervision. This is an asset for medium-sized municipalities or rapidly growing cities where previous traffic records are partial. It also distinguishes the current framework from the majority of data-intensive deep forecasting platforms, which require voluminous archives to maintain performance.[15,36] Another important prerequisite for intelligent transportation systems is the ability to generalize to unseen environments. If traffic topology, demand patterns, or road usage in a single training region differ, models designed exclusively for that region could lose their effectiveness. To assess transfer robustness, the framework was evaluated over unexplored areas that simulated external deployment conditions. As shown in **Figure 15**, the organizational solver retained RMSE reductions of 36.1% with respect to Unseen Region A and 35.3% with respect to Unseen Region B, with associated MAE gains of 35.5% and 31.9%, respectively. Notably, the performance degradation observed during training on unseen regions remains moderate, suggesting that the learned coordination principles have generalized across the spatial range. This implies that the model detects transferable interaction dynamics, rather than memorizing place-specific patterns. This type of behavior is highly relevant to future deployments in cities like Taif, where sensing layouts and traffic behavior differ from those in benchmark datasets. By contrast, rigid centralized models may require lots of refitting when transferred across networks.

The results in **Figures 13–15** support the overall theoretical contribution of this study. The fundamental implication is that traffic intelligence can emerge through distributed computational organization rather than through increasing model depth alone. Reduced latency indicates the efficiency of shallow distributed computation, while the few-shot experiments demonstrate that adaptive coordination improves data efficiency under limited observations. The cross-region experiments further show that the learned organizational coordination principles remain transferable across heterogeneous traffic environments rather than being restricted to fixed spatial layouts.

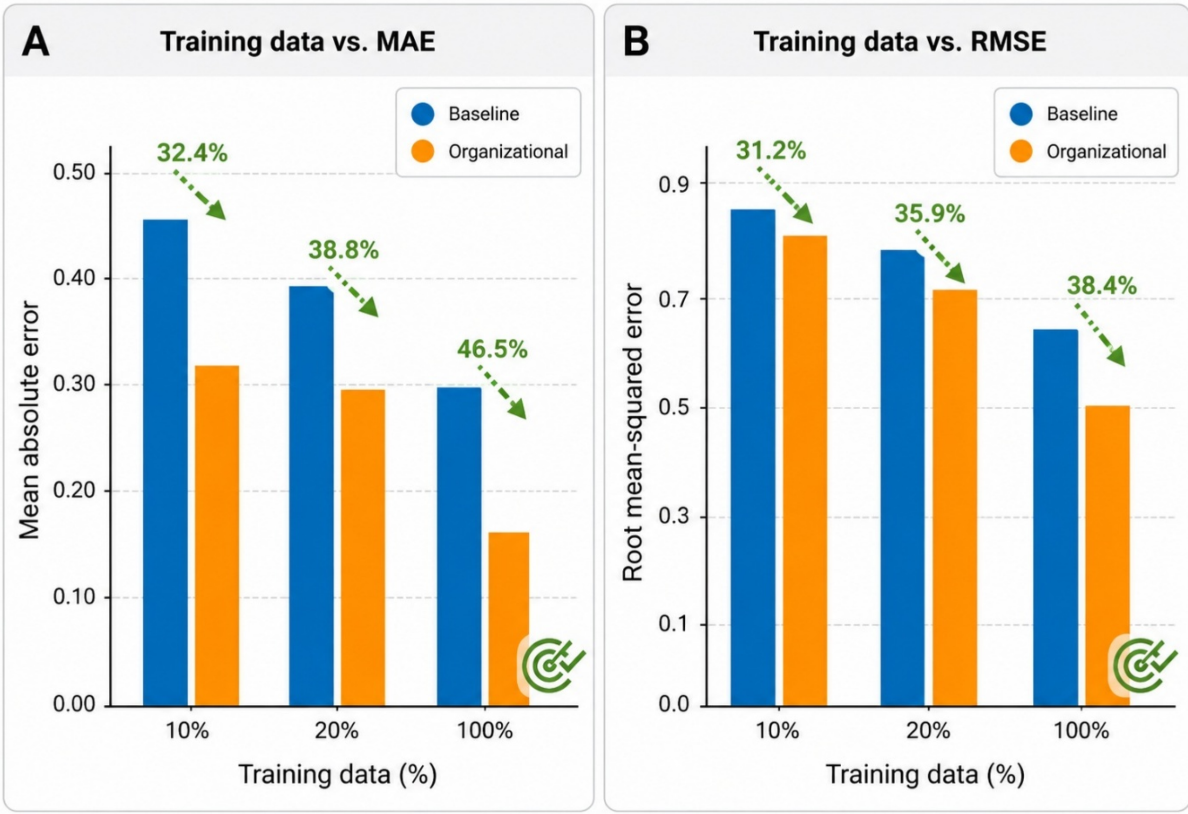

Figure 14. Few-shot performance analysis of the organizational traffic estimation framework. (A) Impact of limited training data on mean absolute error (MAE). (B) Impact of limited training data on root mean-squared error (RMSE).

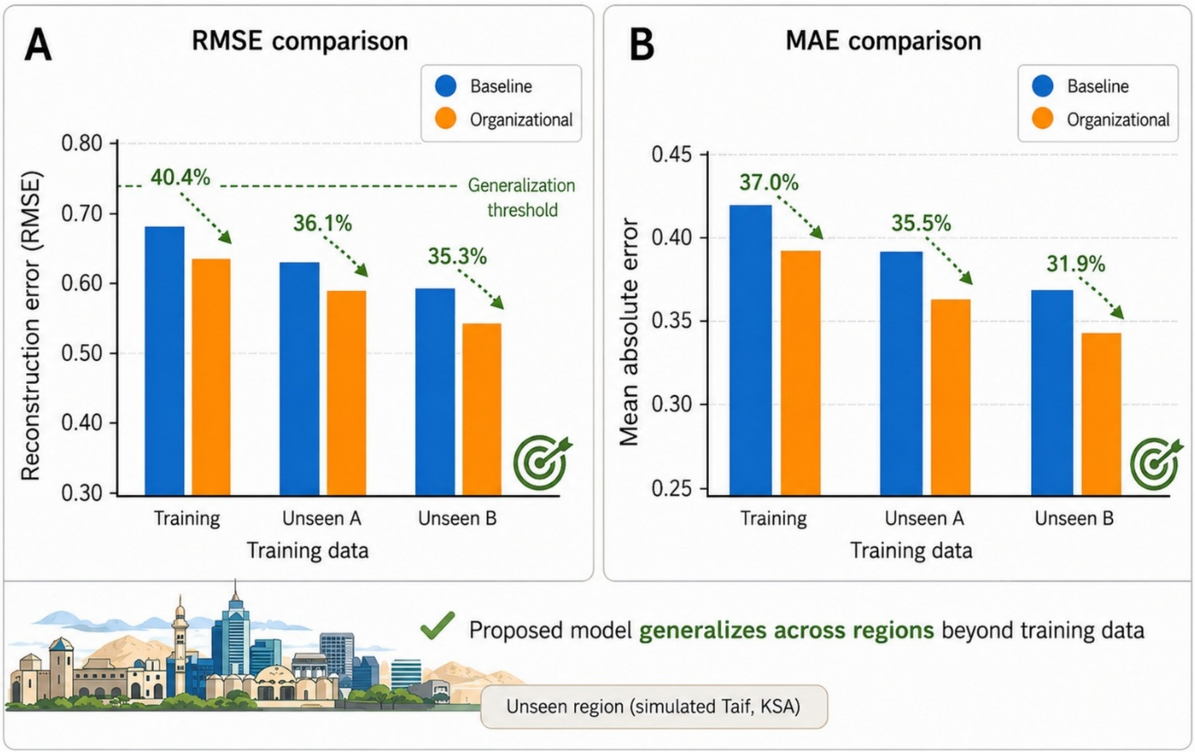

Figure 15. Cross-region transfer and generalization performance of the organizational traffic estimation framework. (A) Root mean-squared error (RMSE) comparison across training and unseen regions. (B) Mean absolute error (MAE) comparison demonstrating robustness in simulated Taif deployment.

Several consistent behaviors were observed across all experimental conditions. Under severe partial observability, the proposed framework consistently achieved lower reconstruction error, improved hidden-node recovery, and stable coordination behavior compared with the centralized baseline. Similar trends were observed across benchmark reconstruction, ablation analysis, few-shot learning, and unseen-region transfer experiments. At the same time, some deployment-scale evaluations showed smaller relative improvements than normalized benchmark experiments because real operational environments contain additional sources of uncertainty, including localized sensor noise, incomplete neighborhood visibility, heterogeneous traffic density, and dynamically changing congestion propagation. These factors increase inference diffusion and reconstruction ambiguity under practical deployment conditions.

More specifically, these findings provide empirical evidence that adaptive structure is a meaningful source of intelligence. In light of recent literature, the results corroborate prior work in multi-agent traffic control, graph forecasting, and distributed decision-support systems.[9–14,25,30,39] However, unlike many existing approaches that rely primarily on deeper centralized architectures or computationally intensive global interactions, the current framework jointly addresses sensing sparsity, computational efficiency, transferability, and distributed adaptation. Overall, the experiments demonstrate that the proposed organizational framework not only improves benchmark reconstruction accuracy but also provides stable, scalable behavior under realistic traffic-estimation conditions with incomplete observations and uncertain traffic dynamics.

6. Conclusion

This study presents a new organizational intelligence model for real-time urban traffic estimation and prediction, treating it as a distributed optimization rather than a conventional centralized learning task. Rather than being grounded in a single deep architecture, the model coordinated diverse, low-weight, compact local computational units that interacted with one another using adaptive influence weighting, role specialization, and structural reorganization. This structure was designed to achieve scalable computation, reliable hidden-state recovery, and adaptive processing for sparse sensing and dynamic traffic conditions. The model clearly outperformed the centralized baseline, as evidenced by experimental measurements. The proposed model reduced the test RMSE by 36.7% and the test MAE by 46.4% on the benchmark dataset, while reducing the validation loss by 48.0%, demonstrating better generalization than the presented model. For real operational units, the framework consistently improved the accuracy of overall and hidden-node reconstruction, demonstrating practical applicability beyond normalized benchmark metrics. Other experiments demonstrated reduced inference latency, enhanced performance with smaller-scale data, and stable transferability to new traffic regions. Such results are consistent with the core proposition of this study: that traffic intelligence may derive meaningfully from enhancing organizational coordination rather than from improved model depth or computational scale alone. Thus, the proposed framework provides a relatively cost-efficient and optimization-based solution to traditional traffic prediction systems, especially in cities where sensing is inadequate and mobility is changing rapidly. This paradigm might be extended in the future into multi-horizon forecasting, signal control optimization, digital twin integration, and uncertainty-aware decision support. On a deeper level, the findings indicate that organizational intelligence is a powerful platform for next-generation intelligent transportation and large-scale distributed optimization systems. Moreover, Future systems may incorporate human expert feedback and interactive DSS interfaces, enabling hybrid intelligence for critical traffic management decisions, and expanding the framework to interact with other smart city subsystems (e.g., energy grids, emergency response systems) can enable holistic urban intelligence and coordinated infrastructure management.

Acknowledgments

The authors would like to express their sincere appreciation to Taif University for providing the research environment and institutional support that facilitated the completion of this work.

Funding

This work is funded by the Deanship of Graduate Studies and Scientific Research, Taif University.

Conflict of interest

The authors declare they have no competing interests or words to that effect.

Author contributions

Conceptualization: Mahmoud Rokaya, El-Sayed Atlam
Data curation: Mahmoud Rokaya

Formal analysis: Mahmoud Rokaya, El-Sayed Atlam
Investigation: Mahmoud Rokaya, Dalia I. Hemdan, Ghada Elmarhomy
Methodology: Mahmoud Rokaya, El-Sayed Atlam
Resources: Ashraf Alyanbaawi, Abdulqader M. Almars, Ghada Elmarhomy
Software: Mahmoud Rokaya
Validation: Mahmoud Rokaya, Ashraf Alyanbaawi, Abdulqader M. Almars
Writing – original draft: Mahmoud Rokaya
Writing – review & editing: El-Sayed Atlam, Dalia I. Hemdan, Ashraf Alyanbaawi, Abdulqader M. Almars, Ghada Elmarhomy

Availability of data

The benchmark dataset used in this study is the publicly available METR-LA traffic dataset (https://doi.org/10.17632/s42kkc5hsw.1), which contains traffic speed measurements collected from urban sensors at 5-minute intervals. The processed data files, experimental configurations, and implementation details supporting the findings of this study are available from the corresponding author upon reasonable request. No proprietary or personally identifiable data were used in this research. Additional implementation resources and related project files are available at: https://taifedusa-my.sharepoint.com/:f:/g/personal/mahmoudrokaya_tu_edu_sa/IgC-pX1SjvUXRYPDWErr4fe5AVhW6TZaCRAlzjHc0KfSViI?e=zYH2q8

AI tools statement

All authors confirm that no AI tools were used in the preparation of this manuscript.

References

1. Sarkhi M. Distributed MADRL with Proactive Traffic Prediction for Scalable and Reliable 6G Resource Allocation. *Int J Intell Eng Syst.* 2025;18(11):629-644. https://doi.org/10.22266/ijies2025.1231.39

2. Shabaz M, Raju KN. AI-Driven Traffic Flow Prediction and Anomaly Detection in Smart Cities: A Multi-Agent Approach. *Trans Emerg Telecommun Technol.* 2025;36(10):e70279. https://doi.org/10.1002/ett.70279

3. Zhao Z, Bi Z, Wang Y, Xie X. A collaborative metaverse-digital twin system for traffic perception, reasoning, and resource scheduling. *Artif Intell Rev.* 2025;59(2):52. https://doi.org/10.1007/s10462-025-11455-9

4. Kosaraju P. AI-based Traffic Prediction and Control. *Int J Artif Intell Data Sci Mach Learn.* 2024;5(4):188-205. https://doi.org/10.63282/3050-9262.IJAIDSML-V5I4P118

5. Shakespear-Miles H, Lin Q, Barzegar S, Ruiz M, Chen X, Velasco L. Centralized and distributed approaches to control optical point-to-multipoint systems near-real-time. *J Opt Commun Netw.* 2024;16(5):565-576. https://doi.org/10.1364/JOCN.516137

6. Liu Q, Tang Y, Li X, Du G, Li Z. Enhancing the Collaborative Decision-Making Performance of Connected and Autonomous Vehicles: A Multi-Modal Failure-Aware Graph Representation Approach. *IEEE Trans Intell Transp Syst.* 2025;26(5):6601-6620. https://doi.org/10.1109/TITS.2025.3534836

7. Zheng B, Rao B, Zhu T, Tan CW, Duan J, Zhou Z, Chen X, Zhang X. Online location planning for AI-defined vehicles: optimizing joint tasks of order serving and spatio-temporal heterogeneous model fine-tuning. *IEEE Trans Mobile Comput.* 2026;25(2):1777-1794. https://doi.org/10.1109/TMC.2025.3608179

8. Wang R, Zhang J, Wang X, He H, Zou Y. Multimodal and multi-agent reinforcement learning framework for urban traffic flow prediction and signal control optimization. *Sci Rep.* 2026;16(1). https://doi.org/10.1038/s41598-026-37722-5

9. Zeynivand A, Javadpour A, Bolouki S, Sangaiah AK, Ja'fari F, Pinto P, Zhang W. Traffic flow control using multi-agent reinforcement learning. *J Netw Comput Appl.* 2022;207:103497. https://doi.org/10.1016/j.jnca.2022.103497

10. Duan M. Attention-based multi-agent reinforcement learning for traffic flow stability in mountainous tunnel entrances. *Sci Rep.* 2025;15(1):37278. https://doi.org/10.1038/s41598-025-21308-8

11. Wu L, Sun Z, Liu J, Shan D, Ma X, Zhu T. Unveiling the impact of heterogeneous driving behaviors on traffic flow: A mesoscale multi-agent modeling approach. *Comput Electr Eng.* 2024;119:109500. https://doi.org/10.1016/j.compeleceng.2024.109500

12. Kušić K, Ivanjko E, Vrbanić F, Gregurić M, Dusparic I. Spatial-temporal traffic flow control on motorways using distributed multi-agent reinforcement learning. *Math.* 2021;9(23):3081. https://doi.org/10.3390/math9233081

13. Mohamed H, Marouane H, Fakhfakh A. A systematic review of multi-agent systems and mobile edge computing in intelligent transportation systems. *J Soft Comput Data Min.* 2025;6(1):169-181. https://doi.org/10.30880/jscdm.2025.06.01.012

14. Chauhan SS, Jain YK, Mannepalli PK, Pandey A. 6G conditioned spatiotemporal graph neural networks for real time traffic flow prediction. *Sci*

Rep. 2026;16(1):3902.
https://doi.org/10.1038/s41598-025-32795-0

15. Zhang X, Li C, Ji L, Kang Y, Pan M, Liu Z, Qi Q. Pretraining-improved Spatiotemporal graph network for the generalization performance enhancement of traffic forecasting. *Sci Rep.* 2025;15(1):27668.
https://doi.org/10.1038/s41598-025-11375-2

16. Mo Z, Liao X, Karbowski DA, Wang Y. Discovering the Precursors of Traffic Breakdowns Using Spatiotemporal Graph Attribution Networks. *arXiv.* Posted online Apr 23, 2025. arXiv:2504.17109.
https://doi.org/10.48550/arXiv.2504.17109

17. Hu J, Liang Y, Fan Z, Liu L, Yin Y, Zimmermann R. Decoupling long-and short-term patterns in spatiotemporal inference. *IEEE Trans Neural Netw Learn Syst.* 2024;35(11):16328-16340.
https://doi.org/10.1109/TNNLS.2023.3293814

18. Anand H, Khayambashi K, Zandsalimi Z, Taghizadeh M, Hasnat MA, Alemazkoor N. Applications of Graph Neural Networks in Civil Infrastructures: A Review on Transportation, Power, Water, and Structural Systems. *SSRN.* Preprint. Posted online Aug 9, 2025.
https://dx.doi.org/10.2139/ssrn.5385353

19. Assolie AA, Khliefat I, Alqatawna A, ALkharabsheh EM, Kasassbeh SA. Machine learning-based prediction of traffic signal timing for optimized intersection management. *Innov Infrastruct Solut.* 2025;10(9):400.
https://doi.org/10.1007/s41062-025-01985-0

20. Chen L, Chen L, Wang H. A Hybrid Spatio-Temporal transformer with Temporal Aggregation and Spatial Memory for traffic flow prediction. *J Big Data.* 2026;13:74.
https://doi.org/10.1186/s40537-026-01383-y

21. Pang Y, Zhang T, Liu Q, Ren J, Huo M, Zhang J. Electric-power telecommunication alarm analysis method based on spatio-temporal graph neural networks. *Int J Inf Commun Technol.* 2025;26(10):1-20.
https://doi.org/10.1504/IJICT.2025.146095

22. Zhang Z, Wan Y, Wang L, Li J. ProMix-DGNet: A Process-Aware Spatiotemporal Network for Sintering System Prediction. *Sensors.* 2026;26(6):1953.
https://doi.org/10.3390/s26061953

23. Deepika MS, Appaji A, Shenoy PD, Venugopal KR. HPC-Accelerated Graph Multi-Scale Spatio-Temporal Transformer for Traffic Forecasting: A Multi-GPU Implementation on METR-LA. In: *2025 Supercomputing India (SCI).* 2025:1-8. IEEE.
https://doi.org/10.1109/SCI68648.2025.113338
47

24. He S, Song R, Chaudhry SS. Service-oriented intelligent group decision support system: application in transportation management. *Inf Syst Front.* 2014;16(5):939-951.
https://doi.org/10.1007/s10796-013-9439-4

25. Khattak AJ, Dahlgren JW, McDonough P. Tools for supporting implementation decisions of intelligent transportation systems. *Transp Res Rec.* 2006;1944(1):41-50.
https://doi.org/10.1177/0361198106194400106

26. He S, Song R, Zhao Q. Design of uncertain group decision support system and its application in intelligent transportation management. In: *Proceedings of the 2003 IEEE International Conference on Intelligent Transportation Systems.* 2003;2:1724-1729.
https://doi.org/10.1109/ITSC.2003.1252778

27. Patel R. Application and future prospects of geographic information system (GIS) in intelligent transportation. *Trans Comput Sci Methods.* 2025;5(3).
https://doi.org/10.5281/zenodo.15070308

28. Balasubramani S, Aravindhar J, Renjith PN, Ramesh K. DDSS: Driver decision support system based on the driver behaviour prediction to avoid accidents in intelligent transport system. *Int J Cogn Comput Eng.* 2024;5:1-13.
https://doi.org/10.1016/j.ijcce.2023.12.001

29. Shulajkovska M, Smerkol M, Noveski G, Bohanec M, Gams M. Artificial intelligence-based decision support system for sustainable urban mobility. *Electronics.* 2024;13(18):3655.
https://doi.org/10.3390/electronics13183655

30. Guido G, Rogano D, Vitale A, Astarita V, Festa D. Big data for public transportation: A DSS framework. In: *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS).* 2017:872-877.
https://doi.org/10.1109/MTITS.2017.8005635

31. Jiang W, Ma Z, Koutsopoulos HN. Deep learning for short-term origin–destination passenger flow prediction under partial observability in urban railway systems. *Neural Comput Appl.* 2022;34(6):4813-4830.
https://doi.org/10.1007/s00521-021-06669-1

32. Othman K, Wang X, Shalaby A, Abdulhai B. Integrating multimodality and partial observability solutions into decentralized multiagent reinforcement learning adaptive traffic signal control. *IEEE Open J Intell Transp Syst.* 2025;6:322-334.
https://doi.org/10.1109/OJITS.2025.3550312

33. Wang X, Taitler A, Smirnov I, Sanner S, Abdulhai B. eMARLIN+: Addressing partial observability to promote traffic signal coordination by leveraging historical information. *IEEE Trans Intell*

Transp Syst. 2024;25(12):21380-21392.
https://doi.org/10.1109/TITS.2024.3462951

34. Tang J, Wan L, Nochta T, Schooling J, Yang T. Exploring resilient observability in traffic-monitoring sensor networks: A study of spatial–temporal vehicle patterns. *ISPRS Int J Geo-Inf.* 2020;9(4):247. https://doi.org/10.3390/ijgi9040247

35. Ghanadbashi S, Golpayegani F. Ontology-enhanced decision-making for autonomous agents in dynamic and partially observable environments. *arXiv.* Posted online May 27, 2024. arXiv:2405.17691. https://doi.org/10.48550/arXiv.2405.17691

36. Maheshwari H, Yang L, Pazzi RW. Machine learning advancements in urban traffic simulation: A comprehensive survey. *IEEE Open J Intell Transp Syst.* 2025;6:1027-1052. https://doi.org/10.1109/OJITS.2025.3589208

37. Chandra J, Navneet SK, Algazinov A, Zhang Y. Safe Urban Traffic Control via Uncertainty-Aware Conformal Prediction and World-Model Reinforcement Learning. *arXiv.* Posted online Feb 2, 2026. arXiv:2602.04821. https://doi.org/10.48550/arXiv.2602.04821

38. Mutambik I. IoT-enabled adaptive traffic management: a multiagent framework for urban mobility optimisation. *Sensors.* 2025;25(13):4126. https://doi.org/10.3390/s25134126

39. Zhang N, Wang FY, Zhu F, Zhao D, Tang S. DynaCAS: Computational experiments and decision support for ITS. *IEEE Intell Syst.* 2008;23(6):19-23. https://doi.org/10.1109/MIS.2008.101

40. Al-Azzawi AA, Zaidan FK, Hassan IA. Designing a control system for traffic lights by VANET protocol. *Int J Nonlinear Anal Appl.* 2022;13(1):1659-1666. https://doi.org/10.22075/ijnaa.2022.5781

41. Wang FY. Parallel control and management for intelligent transportation systems: Concepts, architectures, and applications. *IEEE Trans Intell Transp Syst.* 2010;11(3):630-638. https://doi.org/10.1109/TITS.2010.2060218

42. Taher YH, Mandeep JS, Islam MT, Abdulhae OT, Shakir AT, Islam MS, Soliman MS. Filter for traffic congestion prediction: leveraging traffic control signal actions for dynamic state estimation. *IEEE Access.* 2025;13:8140-8157. https://doi.org/10.1109/ACCESS.2024.3525119

43. Zhao C, Lv Y, Jin J, Tian Y, Wang J, Wang FY. DeCAST in TransVerse for parallel intelligent transportation systems and smart cities: Three decades and beyond. *IEEE Intell Transp Syst Mag.* 2022;14(6):6-17. https://doi.org/10.1109/MITS.2022.3199557

Mahmoud Badee Mahmoud Rokaya *is an Associate Professor of Information Science at Taif University, specializing in optimization, artificial intelligence, and mathematical modeling. He holds a Ph.D. in System Design Engineering from Tokushima University, Japan, with expertise in operations research, ensemble learning, and AI-driven decision support systems. His research integrates deep learning, optimization algorithms, and advanced mathematical models to improve computational efficiency in fields like image processing, water resource management, and AI-based predictive analytics. Dr. Rokaya has led several research projects, published extensively in high-impact journals, and contributed to academic program accreditation and quality assurance. His work in mathematical programming and AI applications enhances smart, data-driven optimization strategies.*

Dalia Ismaeil Ibrahim Hemdan *is a Full Professor of Nutritional Science at Taif University, specializing in therapeutic nutrition, clinical nutrition, and food science. She obtained her Ph.D. in Nutritional Science from Tokushima University, Japan, where she focused on the role of polyphenols in preventing muscle atrophy. Her research encompasses nutritional optimization, bioactive compounds, functional food development, and the impact of antioxidants on disease prevention. She has conducted extensive studies on the therapeutic effects of natural compounds, hepatoprotective agents, and dietary interventions in metabolic disorders and cancer prevention. Dr. Hemdan has been actively involved in academic curriculum development, supervising research projects, and leading funded studies on nutritional interventions, food safety, and dietary optimization. She also explores bibliometric analysis in nutrition science, contributing to a data-driven understanding of dietary impacts on health.*

Ashraf Alyanbaawi *was born in Madinah, Kingdom of Saudi Arabia. He is currently an Assistant Professor at the College of Computer Science and Engineering, Taibah University in Yanbu, KSA. Ashraf received his undergraduate degree from University Tenaga Nasional in Malaysia and began his academic career as a Teaching Assistant at Taibah University. In 2016, he earned both his MS and PhD in Computer Science from Southern Illinois University, USA. His research interests include computer networks, computer security, artificial intelligence, the Internet of Things (IoT),* *and data science.*

Abdulqader Almars *is a Data Scientist and currently serves as an Associate Professor in the Department of Computer Science, College of Science, Northern Border University, Arar, Saudi Arabia. His research interests primarily include data mining, machine learning, data exploration and visualization, social network analytics, and health data analytics.*

Ghada Elmarhomym*: Received her M. Sc. Degrees in Engineering and Ph.D. degree in information science and Intelligent systems from University of Tokushima, Japan, in 2008 and 2012, respectively. She is an assistant professor at the Department of information systems, college of computer science and Engineering, Taibah University, KSA. Her research interests include information retrieval, natural language processing, document processing and Sentence retrieval from huge text data bases, morphological analysis and machine learning.*

El-Sayed Atlam *received his PhD degree in Information Science and Intelligent Systems from Tokushima University, Japan, in 2002. He has been awarded by Japan Society of the Promotion of Science (JSPS) postdoctoral from 2003 to 2005 in same department. He is a professor at the Department of Statistical and Computer science, Tanta University, Egypt. He is a Member in the Computer Algorithm Series of the IEEE and the Egyptian Mathematical Association. He is the Editor Member of the Information Journal of Tokyo and reviewer for many international journals in his field. His research interests include natural language processing, document processing, Data Mining and Machine Learning.*
https://orcid.org/0000-0002-4728-590X

An International Journal of Optimization and Control: Theories & Applications (https://accscience.com/journal/ijocta)