Digital Twin-Enhanced Framework for TCP Throughput Map Construction in Dynamic IoV

Weiwei Hu*, Yuichi Ohsita†, Hideyuki Shimonishi†
*Graduate School of Information Science and Technology, Osaka University, Japan
u865915h@ecs.osaka-u.ac.jp
†D3 Center, Osaka University, Japan
yuichi.ohsita.cmc@osaka-u.ac.jp, shimonishi.cmc@osaka-u.ac.jp

***Abstract*—With the integration of 6G networks and Internet of Vehicles (IoV), constructing reliable Transmission Control Protocol (TCP) throughput maps is crucial for intelligent transportation systems. These maps provide critical insights into network performance, enabling efficient data transmission planning and informed decision-making. However, the dynamic nature of IoV systems, along with resource constraints poses challenges for map construction. To address these challenges, we develop a novel framework for constructing TCP throughput maps. The framework uses base station sensing to collect vehicle positions and communication resources to gather TCP throughput data. A Digital Twin integrates those data over the map in real-time and predicts TCP throughput at every location with prediction uncertainty. A key aspect of this framework is the selection of vehicles to measure TCP throughput, minimizing prediction uncertainty and measurement overhead. We present a method that models the TCP throughput map using Gaussian Process Regression, which accounts for uncertainty. And to take into account the dynamic nature of IoV environments, we also propose the Fixed-Observation Rolling Optimization Algorithm (FOROA) to dynamically select vehicles that measure communication performance. Simulation experiments show that the FOROA reduces map uncertainty 41% faster than the previous method when reducing map uncertainty by 40%.**

***Index Terms*—6G, Digital Twin, IoV, TCP Throughput.**

I. Introduction

As 6G networks and Internet of Vehicles (IoV) technology continue to evolve, constructing a reliable Transmission Control Protocol (TCP) throughput maps has become fundamental for enabling intelligent transportation systems. These maps serve as essential tools for evaluating network performance, managing traffic, and supporting real-time decision-making. In dynamic IoV networks, where communication conditions frequently fluctuate, a TCP throughput map provides valuable guidance for optimizing data routing strategies and enhancing overall system efficiency.

Constructing a TCP throughput map requires two key components: real-time data collection and accurate modeling techniques. Data collection involves acquiring network performance metrics, while modeling integrates this data to represent network conditions coherently and reliably. However, IoV scenarios face challenges such as limited communication resources, high data collection costs, and rapidly changing environments, which complicate both data acquisition and the development of accurate mapping techniques.

To address these challenges, we develop a novel framework for TCP throughput map construction, which dynamically collects TCP throughput data based on vehicle mobility. The key aspect of the framework is the intelligent selection of vehicles for measurement, which plays a critical role in minimizing map uncertainty and optimizing resource utilization.

This framework is based on two kinds of digital twins (DTs), the digital twin of vehicles (DT-V), and the digital twin of communication performance (DT-C). The DT-V serves as a digital representation of vehicle mobility, continuously updated through the collection of vehicle location data from GPS, the analysis of reflective signals from sensing waveforms, as known as Integrated Sensing and Communication (ISAC), and other relevant sources. Leveraging the DT-V, future vehicle locations can be accurately predicted. The DT-C represents a TCP throughput map that incorporates the predicted TCP throughput at each location. It collects TCP throughput data from selected vehicles, using the data to estimate the continuous spatio-temporal TCP throughput map. However, the greater the distance between the observation and estimation points and the longer the time between the observation and estimation times, the less accurate the estimated TCP throughput value will be. Therefore, as proposed in [1], DT-C represents the TCP throughput as a probability map with uncertainty. To capture the spatial and temporal correlations in TCP throughut, we use Gaussian Process Regression (GPR) with a spatio-temporal kernel [2] in DT-C. By using GPR with a spatio-temporal kernel, we can predict the average value of the TCP throught and its uncertainty at each time and location.

We propose a method for selecting vehicles to measure TCP throughput, leveraging the strengths of DT-V and DT-C. This method selects vehicles based on their ability to reduce the uncertainty of the DT-C model, using predicted vehicle locations from DT-V. Traditional optimization and machine learning approaches face significant challenges in IoV networks due to dynamic conditions, high computational demands, and the need for real-time adaptability.

To overcome these limitations, we introduce the **Fixed-Observation Rolling Optimization Algorithm (FOROA)**, tailored for the dynamic nature of IoV environments. FOROA utilizes a fixed observation window to efficiently analyze vehicle trajectories and network conditions, enabling it to

These research results were obtained from the commissioned research (No.JPJ012368C08701) by National Institute of Information and Communications Technology (NICT), Japan.

adapt to rapid changes in mobility and connectivity. By focusing on a limited yet relevant data horizon, FOROA reduces computational overhead while ensuring timely and accurate vehicle selection. This innovation facilitates the construction of precise TCP throughput maps, even under resource and time constraints, making it an ideal solution for IoV applications.

The contributions of this work are summarized as follows:

- **Spatio-Temporal Kernel for TCP Throughput Modeling:** We develop a spatio-temporal kernel function to capture spatial and temporal correlations in TCP throughput data. Integrated within a GPR framework, this kernel enables precise map construction and uncertainty quantification, addressing scenarios with significant spatiotemporal dependencies.
- **Innovative Vehicle Selection with FOROA:** We introduce FOROA, which adapts rolling optimization for IoV scenarios. It efficiently selects vehicles by leveraging observed trajectories and throughput uncertainty, ensuring accurate map construction despite resource constraints.

II. Related Work

The existing research can primarily be classified into two parts: Integration of ISAC with DT technology and Spatio-Temporal-based TCP throughput uncertainty model.

In the development of 6G networks, the integration of ISAC with DT technology is considered a key direction. Recently, several studies have applied DT technology to ISAC systems. In [3], Zhang et al. applied DT technology for sensing channel estimation in cell-free ISAC MIMO systems. In [4], Hu et al. leveraged DT technology in an ISAC system to optimize data processing and reduce long-term computation costs. In the IoV scenario, Ding et al. [5] proposed a DT-based ISAC framework for vehicular networks, employing predictive vehicle tracking and optimized beamforming to enhance transmission rates and sensing accuracy. Gong et al. [6] applied DT technology to an ISAC system within a vehicle edge computing environment, enabling optimized task scheduling and resource allocation by dynamically modeling network conditions and vehicle mobility to reduce response times.

These studies focus on traditional ISAC whose sensing target is real-world objects. On the other hand, our sensing target in this paper is TCP throughput but TCP throughput measurement also consumes radio resources similar to the sensing objects in ISAC. That is, this paper extends the concept of ISAC to include the radio and network environments as the sensing target.

Affected by various factors such as network topology and bandwidth, TCP throughput uncertainty exhibits complex spatio-temporal dependencies. Hoang et al. investigated the spatial correlation of TCP throughput by analyzing the joint impact of transmission errors in satellite-to-vehicle last-mile channels and congestion losses on the Internet [7]. Bommisetty et al. analyzed the temporal correlation of TCP throughput by studying how Round-Trip Time and buffer size affect queue metrics, revealing the dependency of throughput stability on varying network conditions [8].

While these studies provide valuable insights into the spatial and temporal correlation of TCP throughput, they often lack a unified approach to model the joint spatio-temporal dependencies of TCP throughput uncertainty, especially in dynamic and complex IoV scenarios. To address this gap, our work introduces a spatio-temporal approach to model TCP throughput uncertainty, effectively capturing its variations for dynamic tracking and prediction in IoV scenarios.

III. System model

A. The Introduction of our System

As shown in Figure 1, We consider a DT-Enhanced vehicle selection Framework comprising a physical environment and a DT environment. The physical environment includes several uniformly distributed observation points $\mathcal{G} = \{g_1, \ldots, g_n, \ldots, g_N\}$ and several vehicles participating in the collection of TCP throughput data $\mathcal{V} = \{v_1, \ldots, v_m, \ldots, v_M\}$. The DT environment is maintained by a base station and consists of two components: DT-V and DT-C. The base station provides the DT environment with vehicle location data and TCP throughput data collected by vehicles. The DT-V processes vehicle location data to predict future vehicle trajectories, while the DT-C integrates throughput data to update the TCP throughput map and compute uncertainty values. In return, the DT-V provides vehicles with predicted trajectory data, and the DT-C provides throughput uncertainty values, assisting the base station in selecting vehicles for data collection. The process for updating DT-V and DT-C is shown as follows:

Initially, the base station collects real-time vehicle position data via GPS and other onboard sensors. This data is used to update the DT-V, which then predicts future vehicle locations. Simultaneously, the DT-C provides the base station with the uncertainty of TCP throughput from the previous time step. Combining the predicted vehicle locations from DT-V and the TCP throughput uncertainty of the previous time step from DT-C, the base station selects specific vehicles to monitor TCP throughput. The selected vehicles measure throughput data and transmit it back to the base station. This data is integrated into the DT-C to update the TCP throughput map and compute its associated uncertainty, further improving the model's accuracy.

In the rest of this paper, we focus on the DT-C; we introduce the spatio-temporal model used in the DT-C and the method to select vehicles for TCP throughput measurement, assuming that the DT-V can predict the future locations of vehicles accurately.

B. Spatio-Temporal Correlation Model

The construction of a TCP throughput map in this framework involves two essential components: (1) predicting mean throughput values for observation points and (2) quantifying and updating the uncertainty associated with these predictions. To achieve this, we employ GPR with a custom-designed spatio-temporal kernel function, which captures the dependencies in TCP throughput data and enables precise prediction and uncertainty modeling.

The spatio-temporal kernel function serves as the backbone of the GPR framework, modeling the relationships between

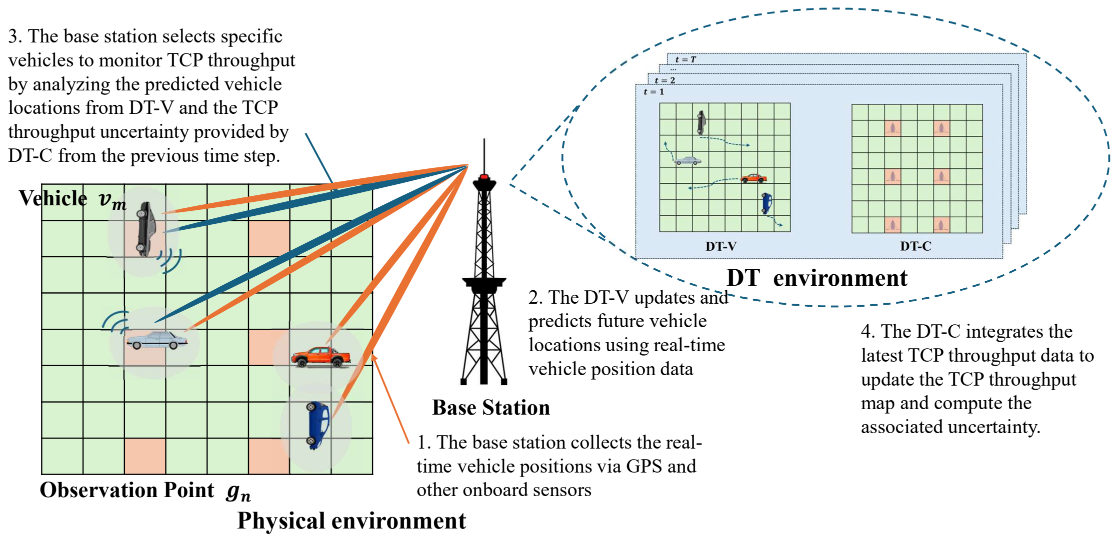

Fig. 1. Digital Twin-Enhanced vehicle selection Framework

vehicles and observation points by capturing both spatial and temporal correlations. This kernel function is defined as:

$$k_{ST}(v_m, g_n) = \beta^2 \cdot \exp\left(-\frac{\|z_m - z_n\|^2}{2l_s^2}\right) \cdot \exp\left(-\frac{|t_m - t_n|}{2l_t}\right) \quad (1)$$

where β^2 is the process variance parameter, representing the variability of the throughput data. $\exp\left(-\frac{\|z_m - z_n\|^2}{2l_s^2}\right)$ captures spatial correlations, with l_s being the spatial length scale and z_m, z_n denoting the positions of v_m and g_n, respectively. $\exp\left(-\frac{|t_m - t_n|}{2l_t}\right)$ captures temporal correlations, where t_m is the time step of v_m, and t_n is the time step of g_n. This kernel function enables the GPR model to account for dynamic spatio-temporal dependencies, providing a robust framework for constructing the TCP throughput map.

Using the spatio-temporal kernel function, GPR predicts the mean TCP throughput value $\mu(g_n^t)$ for each observation point g_n at time step t as follows:

$$\mu(g_n^t) = \sum_{v_m \in \mathcal{S}^t} w(v_m, g_n) \cdot T(v_m) \quad (2)$$

where $T(v_m)$ represents the observed throughput data collected from vehicle v_m; $w(v_m, g_n) = \frac{k_{ST}(v_m, g_n)}{\sum_{v'_m \in \mathcal{S}^t} k_{ST}(v'_m, g_n)}$ is the normalized weight derived from the spatio-temporal kernel, indicating the influence of v_m on g_n; $\mathcal{S}^t \subseteq \mathcal{V}^t$ is the subset of selected vehicles at time step t. This prediction forms the foundation of the TCP throughput map, representing mean throughput values across the road network.

In addition to predicting the mean values, GPR quantifies the uncertainty $\sigma(g_n^t)$ associated with each observation point g_n. The uncertainty is iteratively updated at every time step t using the following rule:

$$\sigma(g_n^t) = \max\left(0, \sigma(g_n^{t-1}) - k_{ST}(g_n^t, \mathcal{S}^t)\right) \quad (3)$$

where $\sigma(g_n^{t-1})$ is the uncertainty from the previous time step; $k_{ST}(g_n^t, \mathcal{S}^t)$ is the aggregated spatio-temporal kernel value between g_n and the selected vehicle set $\mathcal{S}^t$; $\max(0, \cdot)$ ensures non-negative uncertainty values.

By iteratively refining the uncertainty values, the framework achieves a more precise representation of the confidence levels associated with each throughput prediction.

The total uncertainty across the road network at time step t is computed as:

$$U(t) = \sum_{g_n \in \mathcal{G}} \sigma(g_n^t) \quad (4)$$

The accumulated uncertainty over $[t, t']$ is defined as:

$$\mathcal{U} = \int_t^{t'} U(t)\, dt = \int_t^{t'} \sum_{g_n \in \mathcal{G}} \sigma(g_n^t)\, dt \quad (5)$$

By leveraging the spatio-temporal kernel function and the GPR, our framework constructs a TCP throughput map that integrates predicted throughput values with associated uncertainties. This enables precise modeling and supports decision-making in dynamic IoV environments.

IV. Optimizing the Uncertainty of TCP Throughput Map with FOROA

In this section, we present the methodology for minimizing the uncertainty of the TCP throughput map using the FOROA. The mathematical formulation and integration of the core elements—state, action, and reward—are detailed below.

A. Problem Formulation

To explore the best vehicle selection strategy and collect TCP throughput data, the optimization problem is defined as:

$$\begin{aligned} \min_{\mathcal{S}^t} & \int_t^{t'} \sum_{g_n \in \mathcal{G}} \sigma(g_n^t)\, dt \\ \text{s.t.} & \sum_{v_m^t \in \mathcal{S}^t} \mathbb{1}(v_m^t) \le K^t, \quad \forall t, \\ & \mathcal{S}^t \subseteq \mathcal{V}^t, \quad g_n \in \mathcal{G} \end{aligned} \quad (6)$$

where $\mathbb{1}(v_m^t)$ is an indicator function returning 1 if vehicle v_m is selected, and 0 otherwise; K^t is the maximum number of vehicles that can be selected at time t; $\mathcal{G}$ represents the set of all observation points in the road map; and $\sigma(g_n^t)$ is the uncertainty at observation point g_n at time t.

This problem seeks to minimize the total accumulated uncertainty over the observation horizon by optimally selecting vehicles. However, due to the non-linear dependencies introduced by the spatio-temporal kernel $k_{ST}(g_n^t, \mathcal{S}^t)$, the optimization problem is non-convex, posing significant challenges for traditional optimization techniques.

B. Dynamic Vehicle Selection Using FOROA

To tackle the challenge of dynamic vehicle selection in IoV environments, we propose FOROA, an algorithm designed to minimize the uncertainty of the TCP throughput map while effectively operating within limited resources. FOROA operates as a feedback loop and progresses through three stages: (1) **Observation and Assessment**: During this phase, the algorithm uses the state s^t to observe vehicle movements over a fixed observation window c and assess the current uncertainty levels in the TCP throughput map. (2) **Evaluation**: In this stage, the algorithm calculates the reward r^t, which quantifies the reduction in uncertainty achieved by each vehicle's data contribution. This reward is used to evaluate each vehicle's potential to improve the TCP throughput map. (3) **Selection**: Finally, the algorithm determines the action a^t, which corresponds to selecting an optimal subset of vehicles. This subset is chosen to maximize the reduction in uncertainty across the TCP throughput map while adhering to resource constraints. These stages are supported by the definitions of state, action, and reward, as detailed below.

1) State: The *state* s^t captures all relevant information available to the base station at time t, enabling the algorithm to make informed decisions. The state includes the predicted future positions of vehicles, the current observation point, and the uncertainty levels from the previous time step:

$$s^t = \left(\{z_m^{t+1,\dots,t+c}\}, z_n^t, \{\sigma(g_i^{t-1}) \,|\, i = 1, \dots, N\}, \mathcal{V}^t\right) \quad (7)$$

2) Action: The *action* a^t corresponds to selecting a subset of vehicles $\mathcal{S}^t \subseteq \mathcal{V}^t$ for data collection at time t. The algorithm evaluates the contribution of each vehicle within the observation window and selects vehicles expected to most effectively reduce the uncertainty of the TCP throughput map. Formally, the action is defined as:

$$a^t = \mathcal{S}^t \quad (8)$$

3) Reward: The *reward* r^t quantifies the effectiveness of the selected vehicles in improving the TCP throughput map. It is calculated as the total reduction in uncertainty across all observation points:

$$r^t = \sum_{g_n \in \mathcal{G}} \left(\sigma(g_n^{t-1}) - \sigma(g_n^t)\right) \quad (9)$$

By integrating state, action, and reward, FOROA ensures robust and adaptive vehicle selection, enabling efficient and accurate updates to the TCP throughput map in dynamic IoV environments.

V. Simulation Evaluation

A. Simulation Settings

The experiment is conducted using **OMNeT++**, **SUMO**, and **Veins** to simulate vehicle mobility and data collection in a complex IoV scenario. First, **SUMO** is used to construct the road model, followed by **OMNeT++** and **Veins** for building the vehicle mobility model and the road TCP throughput distribution model. The objective is to select vehicles in real time to collect TCP throughput data and minimize the uncertainty $\sigma(g_n^t)$ of the TCP throughput map. Table I summarizes the simulation settings used in the experiment.

TABLE I
Simulation Settings

Parameters	**Values**
The number of observation points N	304
Uncertainty of observation point $\sigma(g_n^t)$	0-1
Maximum Number of Vehicles M	20.0
Time Steps T	0-99
Length Scale (Space) l_s	60.0
Length Scale (Time) l_t	5.0
Maximum Vehicles to Select (per step) K^t	1-5
Process Variance β^2	1.0
Window Size c	0-5

B. Convergence Analysis of FOROA algorithm

We investigate the impact of window size and the number of vehicles selected on the convergence of the FOROA algorithm.

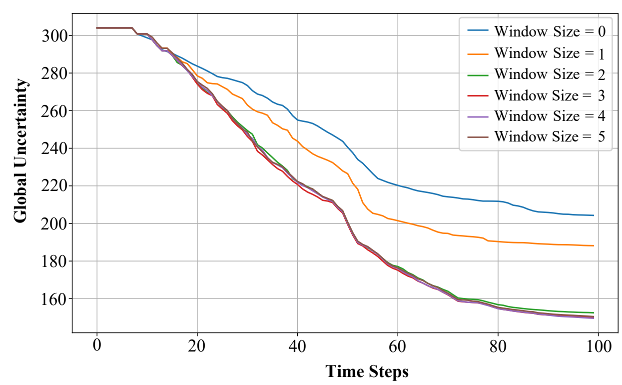

Fig. 2. Window Size Impact on the Uncertainty

Figure 2 illustrates the impact of varying the window size parameter, which predicts future vehicle positions, on the uncertainty of the TCP throughput map. Experimental results indicate that increasing the window size beyond 2 does not significantly reduce uncertainty. This suggests diminishing returns in uncertainty reduction with larger window sizes, likely due to the reduced relevance of distant future predictions to current decisions. Consequently, we fix the window size at 2 in our algorithm, balancing computational efficiency with minimizing parameter size to ensure optimal performance without unnecessary computational overhead.

Figure 3 illustrates how the number of vehicles selected per time step affects the reduction of uncertainty in the TCP

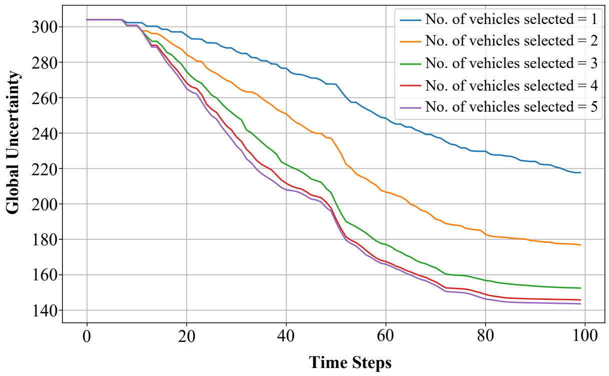

Fig. 3. Number of Vehicle Slected Impact on the Uncertainty

throughput map, with the window size fixed at 2. The experimental results show that uncertainty decreases significantly as the number of selected vehicles increases from 1 to 3. However, beyond 3 vehicles, additional selections provide diminishing returns in uncertainty reduction. This trend indicates that selecting more than 3 vehicles does not substantially enhance the reduction of uncertainty in the TCP throughput map. Therefore, setting the maximum number of vehicles selectable per second to 3 strikes an efficient balance between cost and performance, enabling effective and economical road TCP throughput data collection.

C. Comparison with other algorithms

In this section, we fix the window size to 2 and limit the maximum number of selected vehicles to 3 per time step. The performance of FOROA is compared with three methods: **Select All Vehicles (SAV)**, which selects all available vehicles at each time step to achieve the fastest uncertainty reduction; **Greedy Selection Algorithm (GSA)**, which sequentially selects three vehicles by prioritizing the immediate reduction of TCP throughput uncertainty at each time step, without considering the overall global optimization, aims to minimize uncertainty; and **Random Selection Algorithm (RSA)**, which randomly selects up to three vehicles per time step.

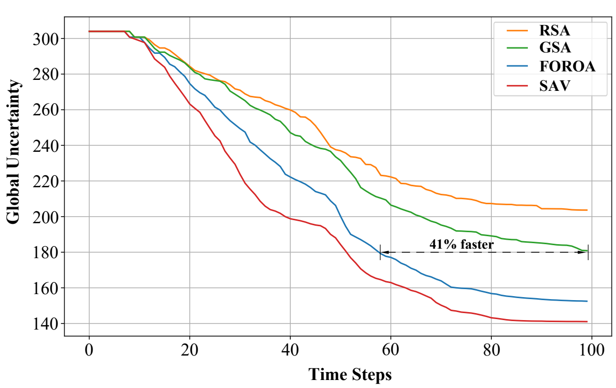

Fig. 4. Comparison of FOROA with Three Other Algorithms

As illustrated in Fig. 4, the FOROA algorithm demonstrates superior performance compared to RSA by effectively utilizing resources, significantly reducing uncertainty, and adapting to dynamic IoV environments. Specifically, when reducing map uncertainty by 40%, FOROA achieves this 41% faster than GSA, leveraging its predictive capabilities to optimize resource allocation more effectively under limited resources for TCP throughput measurement. In addition, compared to SAV, FOROA achieves near-optimal performance despite constraints on the number of selected vehicles and limited observation of vehicle trajectory time steps. These results underscore the algorithm's efficiency in resource utilization, robustness against temporal constraints, and its ability to maintain low uncertainty in the TCP throughput map.

VI. Conclusion

In this work, we develop a novel framework for reducing uncertainty in the TCP throughput map within dynamic IoV networks, leveraging DT technologies. To address resource constraints and dynamic environmental challenges, we propose the FOROA, a heuristic approach that optimizes vehicle selection within a fixed observation window size. By integrating spatio-temporal kernel functions with a rolling optimization strategy, FOROA achieves efficient and cost-effective solutions for real-time TCP throughput data collection and uncertainty reduction. This framework lays a solid foundation for constructing scalable TCP throughput maps in dynamic IoV scenarios.

Future research will focus on visualizing the TCP throughput map to provide intuitive insights and exploring more advanced predictive models to further enhance the framework's accuracy and applicability.

References

[1] H. Shimonishi, D. Kominami, Y. Ohsita, M. Masayuki, and et al., "Probabilistic representation and its application of digital-twin of spatio-temporal real-world toward trustable cyber-physical interactions," *IEEE Network*, vol. 38, no. 6, pp. 130–137, 2024.

[2] T. Gao, T. Nattaon, Y. Ohsit, and H. Shimonishi, "Robot path planning for monitoring dynamic environment by predictive uncertainty minimization using gaussian process regression," *IEEE 43rd International Conference on Consumer Electronics (ICCE)*, 2025.

[3] J. Zhang, S. Xu, Z. Zhu, R. Yang, C. Li, Y. Huang, and L. Yang, "Digital-twin-enabled sensing channel estimation for 6g cell-free isac mimo system," in *2024 IEEE Wireless Communications and Networking Conference (WCNC)*, 2024, pp. 1–6.

[4] S. Hu, J. Gao, X. Huang, M. Li, K. Qu, C. Zhou, and X. Shen, "Digital twin-based user-centric edge continual learning in integrated sensing and communication," in *ICC 2024 - IEEE International Conference on Communications*, 2024, pp. 5646–5651.

[5] W. Ding, Z. Yang, M. Chen, Y. Liu, and M. Shikh-Bahaei, "Joint vehicle connection and beamforming optimiziation in digital-twin-assisted integrated sensing and communication vehicular networks," *IEEE Internet of Things Journal*, vol. 11, no. 20, pp. 32 923–32 938, 2024.

[6] Y. Gong, Y. Wei, Z. Feng, F. R. Yu, and Y. Zhang, "Resource allocation for integrated sensing and communication in digital twin enabled internet of vehicles," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 4, pp. 4510–4524, 2023.

[7] H. D. Le, P. V. Trinh, and et al., "Throughput analysis for tcp over the fso-based satellite-assisted internet of vehicles," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 2, pp. 1875–1890, 2022.

[8] L. Bommisetty, "Performance analysis of tcp queues: Effect of buffer size and round trip time," in *2021 6th International Conference on Signal Processing, Computing and Control (ISPCC)*, 2021, pp. 631–635.