xURLLC-Aware Service Provisioning in Vehicular Networks: A Semantic Communication Perspective

Le Xia, *Graduate Student Member, IEEE*, Yao Sun, *Senior Member, IEEE*, Dusit Niyato, *Fellow, IEEE*, Daquan Feng, *Member, IEEE*, Lei Feng, *Member, IEEE*, and Muhammad Ali Imran, *Fellow, IEEE*

***Abstract*—Semantic communication (SemCom), as an emerging paradigm focusing on meaning delivery, has recently been considered a promising solution for the inevitable crisis of scarce communication resources. This trend stimulates us to explore the potential of applying SemCom to wireless vehicular networks, which normally consume a tremendous amount of resources to meet stringent reliability and latency requirements. Unfortunately, the unique background knowledge matching mechanism in SemCom makes it challenging to simultaneously realize efficient service provisioning for multiple users in vehicle-to-vehicle networks. To this end, this paper identifies and jointly addresses two fundamental problems of knowledge base construction (KBC) and vehicle service pairing (VSP) inherently existing in SemCom-enabled vehicular networks in alignment with the next-generation ultra-reliable and low-latency communication (xURLLC) requirements. Concretely, we first derive the knowledge matching based queuing latency specific for semantic data packets, and then formulate a latency-minimization problem subject to several KBC and VSP related reliability constraints. Afterward, a SemCom-empowered Service Supplying Solution (S^4) is proposed along with the theoretical analysis of its optimality guarantee and computational complexity. Numerical results demonstrate the superiority of S^4 in terms of average queuing latency, semantic data packet throughput, user knowledge matching degree and knowledge preference satisfaction compared with two benchmarks.**

***Index Terms*—Semantic communication, vehicular networks, next-generation ultra-reliable and low-latency communication, knowledge base construction, vehicle service pairing.**

I. Introduction

UBIQUITOUS intelligence is expected to emerge in next-generation vehicular networks to accommodate diverse smart on-board applications and large-capacity vehicle-to-everything (V2X) services (e.g., multimodal artificial intelligence-generated content offered by ChatGPT or Dall-E), which poses tremendous demands on high data rates along with stringent requirements for reliability and latency [2], [3]. Correspondingly, the scarcity of available communication resources, such as bandwidth and energy, is envisioned to exacerbate to unprecedented levels and to be the most challenging problem in the near future, especially considering traditional communications-based massive vehicular networks. Fortunately, *semantic communication* (SemCom) beyond the conventional Shannon paradigm has recently been recognized as a promising remedy for communication resource saving and transmission reliability promotion [4]–[11], which, therefore, inspires us to investigate the potential of exploiting SemCom to perform efficient service provisioning in vehicular networks.

Different from traditional communication schemes of guaranteeing the precise reception of transmitted bits [12], the accurate delivery of semantics implied in desired messages becomes the cornerstone concept of SemCom [4]. Taking a single SemCom-enabled vehicle-to-vehicle (V2V) link as an example, a sender vehicle first leverages background knowledge relevant to source messages to filter out irrelevant content and extract core semantic features that only require fewer bits for transmission, the process of which is called semantic encoding [5]. Once the receiver vehicle has the same knowledge as the sender, its local semantic interpreters are capable of accurately restoring the original meanings from the received bits, even with intolerable bit errors during data dissemination [7]. This process is called semantic decoding [5]. As a matter of fact, the rationale behind the growing prosperity of SemCom is due to the increasingly powerful representation and generalization abilities of semantic coding models. Particularly, powered by state-of-the-art deep learning (DL) algorithms-based semantic models, SemCom has been well explored to provision multimodal services, including text [6], image [8], video [11], and so on. Consequently, SemCom is believed to significantly alleviate the resource scarcity problem, while ensuring sufficient transmission efficiency along with ultra-low semantic errors in V2V networks.

Apart from many superiorities, it is worth pointing out that equivalent background knowledge should be of paramount importance to eliminate semantic ambiguity, which has led to a key concept of *knowledge base* (KB) in the realm of SemCom [7]–[11]. Specifically, a single KB is deemed a small information entity that contains background knowledge corresponding to only one particular application domain (e.g.,

Preliminary results of this work have been presented in part at the IEEE International Conference on Communications (ICC), 2023 [1]. (*Corresponding author: Yao Sun.*)

Le Xia, Yao Sun, and Muhammad Ali Imran are with the James Watt School of Engineering, University of Glasgow, Glasgow G12 8QQ, UK (e-mail: l.xia.2@research.gla.ac.uk; {Yao.Sun, Muhammad.Imran}@glasgow.ac.uk).

Dusit Niyato is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (e-mail: dniyato@ntu.edu.sg).

Daquan Feng is with the Guangdong Province Engineering Laboratory for Digital Creative Technology, Shenzhen University, Shenzhen 518060, China (e-mail: fdquan@szu.edu.cn).

Lei Feng is with the School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: fenglei@bupt.edu.cn).

music or sports) [7].[1] Since different KBs are associated with different background knowledge, holding some common KBs becomes the necessary condition to perform SemCom between two vehicles in accordance with the knowledge equivalence principle. Moreover, through employing differing KBs, semantic information related to different application domains can be accurately exchanged among vehicles, thereby efficiently achieving service provisioning in a way of SemCom.

By reviewing the recent research advancement of SemCom, there have been some preliminary publications addressing several challenges of semantics-aware wireless networks. Xie *et al.* [6] devised a Transformer-based end-to-end SemCom system to realize link-level text transmission, and then improved this system to be lightweight in [14]. Proceeding as in [6], Yan *et al.* [15] exploited the semantic spectral efficiency optimization-based channel assignment. In parallel, Xia *et al.* [9] developed the bit-to-message transformation and a new metric called system throughput in message for the first time to optimize resource management in SemCom-enabled cellular networks. Nevertheless, to the best of our knowledge, none of the existing work has ever explored the potential of applying SemCom to V2V networks in alignment with stringent latency and reliability requirements, which should be rather challenging as explained below.

In full view of the novel paradigm of *SemCom-enabled vehicular networks* (SCVNs), the task lies in seeking the optimal solution to efficiently provide all participating vehicle users (VUEs) with diverse SemCom-empowered services. However, it is noticed that enabling the next-generation ultra-reliable and low-latency communication (xURLLC) remains indispensable, especially when pursuing adequate semantic fidelity for large-scale V2V communications. Uniquely, the reliability requirement originates from the aforementioned strict knowledge equivalence condition, while the latency requirement is related to varying processing efficiencies of semantic interpretation models. In summary, we are encountering three fundamental networking challenges in SCVN.

- *Challenge 1: How to measure performance in terms of reliability and latency when introducing SemCom into vehicular networks?* Notice that data packets transmitted in traditional V2V communications generally consider only one type of queuing process, in which different packets have the same distributions of arrival and interpretation [16]. However, semantic data packets in SCVN related to various SemCom-empowered services may result in different queuing and processing delays due to the different semantic interpretation models equipped on VUEs. Hence, it is not trivial to accurately measure and assume prior information about the latency performance in SCVN. Besides, given the core mechanism of semantic delivery, it should be more reasonable to characterize the reliability performance from the knowledge equivalence perspective between any two associated VUEs for V2V communications. All of the above constitutes the first and the main challenge.
- *Challenge 2: How to construct appropriate KBs at each VUE for better SemCom-empowered service provisioning?* Considering varying practical KB sizes, personal preferences on different SemCom services, and the limited vehicular storage capacities, there is a pressing need to devise an optimal *knowledge base construction* (KBC) policy that is not only proactive but also collaborative for all VUEs to construct their respective appropriate KBs for better service provisioning. Notably, this challenge is inevitable in xURLLC-aware SCVN, since the remote KB access approach (i.e., the approach that each VUE remotely accesses its required KBs via RSUs, Cloud servers or core networks) can incur intolerable communication overhead and transmission latency. Therefore, the local and distributed KBC approach should be more applicable for each VUE to well perform SemCom.
- *Challenge 3: How to select the best vehicle node for each VUE from multiple candidate neighbors to optimize service provisioning related overall network performance?* As mentioned earlier, selecting vehicle pairs for realizing service provisioning is very much distinct in SCVN due to the knowledge equivalence condition. Combined with different KBs constructed at numerous VUEs and unstable wireless link quality, it can be challenging to well solve the service provisioning-driven vehicle pairing problem, namely *vehicle service pairing* (VSP).

In line with the above, it is particularly worthwhile to note that challenges 2 and 3 are closely coupled, which makes it indispensable to jointly seek the optimal KBC and VSP policy for all VUEs to meet the xURLLC requirements. Moreover, efficient SemCom-empowered service provisioning is excepted after addressing the three challenges to yield a bunch of benefits in SCVN, such as improving V2V information interaction efficiency, reducing data traffic congestion, and ensuring high-quality vehicular services.

In this paper, we propose a novel SemCom-empowered Service Supplying Solution (S^4) in SCVN with the awareness of meeting the xURLLC requirements. Both theoretical analysis and numerical results demonstrate the performance superiority of S^4 in terms of average queuing latency, semantic data packet throughput, user knowledge matching degree, and user knowledge preference satisfaction compared with two different benchmarks. In a nutshell, our main contributions are summarized as follows:

- We identify two fundamental yet unique problems KBC and VSP in SCVN by fully incorporating SemCom-related characteristics with vehicular network scenarios. In particular, individual VUE preference for different KBs is considered in the KBC, while the VSP of two adjacent VUEs takes into account the strict matching requirement between their respective constructed KBs.
- We theoretically derive the KB matching based queuing latency for a VUE pair in SCVN. Then, through carefully analyzing the unique queuing features of received semantic data packets, a joint latency-minimization problem

[1]The structure of a KB can roughly cover multiple computational ontologies, facts, rules and constraints associated to a specific domain [13]. In recent DL-driven semantic coding models, the KB is also treated as a training database serving a certain class of learning tasks [7], [10]. Relevant research is beyond the scope of this paper and will not be discussed in depth.

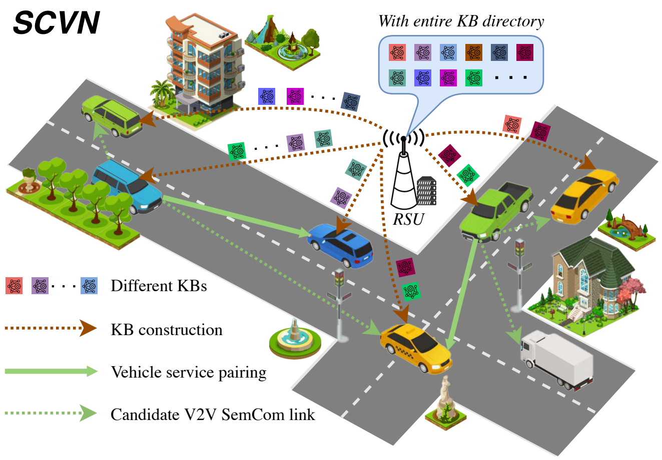

Fig. 1. The SCVN scenario with KBC and VSP.

is mathematically formulated subject to several KBC and VSP-related reliability constraints and other practical system limitations.

- We develop an efficient solution named S^4 to tackle the above optimization problem, and its optimality is theoretically proved by two propositions. Specifically, a primal-dual problem transformation method is first exploited in S^4 to obtain the corresponding Lagrange dual problem, followed by a two-stage method dedicated to solving multiple subproblems with a low computational complexity. Given the dual variables in each iteration, the first stage is to obtain the optimal KBC sub-policy for each potential VUE pair, whereby the second stage is able to finalize the optimal solutions of KBC and VSP for all VUEs in SCVN.

The remainder of this paper is organized as follows. Section II first introduces the system model of SCVN. Then a joint service provisioning problem in SCVN is identified and formulated in Section III. In Section IV, we illustrate the proposed solution S^4. Numerical results are presented and discussed in Section V, followed by conclusions in Section VI.

II. System Model

In this section, the considered SCVN scenario is first elaborated along with the knowledge storage model and vehicle pairing model. Then, the knowledge matching based queuing latency for semantic packets is derived.

A. SCVN Scenario

Consider an SCVN scenario as shown in Fig. 1, the total of V VUEs are distributed within the coverage of a single roadside unit (RSU), and each VUE $i \in \mathcal{V} = \{1, 2, \ldots, V\}$ is capable of providing SemCom-empowered services to others. For the wireless propagation model, let $\gamma_{i,j}$ denote the signal-to-interference-plus-noise ratio (SINR) experienced by the V2V link between VUE i and VUE j ($j \neq i$). Essentially, we allow VUE i to communicate with VUE j if their SINR value $\gamma_{i,j}$ is above a prescribed threshold γ_0. In this manner, the set of communication neighbors of VUE i is defined as $\mathcal{V}_i = \{j | j \in \mathcal{V}, j \neq i, \gamma_{i,j} \geqslant \gamma_0\}$, $\forall i \in \mathcal{V}$. Moreover, it is known that the RSU has powerful communication, computing, and storage capabilities and can provide stable communication coverage [2]. Hence, in this work, let the RSU act as a semantic service controller in the SCVN to efficiently schedule and coordinate the whole SemCom-empowered service provisioning process based on the request and state information received from all participating VUEs within its coverage.

B. Vehicular Knowledge Storage Model

Due to the unique mechanism of semantic interpretation, the acquisition of necessary background knowledge is inevitable for all SemCom-enabled transceivers. In this work, assuming that all VUEs are able to proactively download and construct their respective required KBs from the RSU, where each VUE i has a finite capacity C_i for its local KB storage. Meanwhile, suppose that there is a KB library $\mathcal{K}$ with a total of N differing KBs in the considered SCVN, and each requires a unique storage size s_n, $n \in \mathcal{K} = \{1, 2, \ldots, N\}$. Furthermore, we define a binary KBC indicator as

$$\alpha_i^n = \begin{cases} 1, & \text{if KB } n \text{ is constructed at VUE } i; \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

It is worth mentioning that the same KB cannot be constructed repeatedly at one VUE for reducing redundancy and for promoting the storage efficiency.

Besides, it is noticed that different VUEs may have different preferences for these KBs corresponding to their required services, thus resulting in the diversity of KB popularity. Naturally, the more popular the KBs, the higher the KBC probabilities. Therefore, without loss of generality, we assume that the KB popularity at each VUE follows Zipf distribution [17].[2] Hence, the probability of VUE i requesting its desired KB n-based services (generating the corresponding semantic data packets) is $p_i^n = (r_i^n)^{-\xi_i} / \sum_{e \in \mathcal{K}} e^{-\xi_i}, \forall (i,n) \in \mathcal{V} \times \mathcal{K}$, where ξ_i ($\xi_i \geqslant 0$) is the skewness of the Zipf distribution, and r_i^n is the popularity rank of KB n at VUE i.[3] Based on p_i^n, we specially develop a KBC-related metric η_i, namely *knowledge preference satisfaction*, to measure the satisfaction degree of VUE i constructing its interested KBs as

$$\eta_i = \sum_{n \in \mathcal{K}} \alpha_i^n p_i^n. \tag{2}$$

It is further required that $\eta_i \geqslant \eta_0$, where η_0 is the unified minimum threshold that needs to be achieved at each VUE.

C. Vehicle Pairing Model for SemCom

Apart from equipping with suitable KBs, it is also essential for each VUE to select an appropriate VUE from its neighbors for SemCom-empowered service provisioning. It is worthwhile to re-emphasize that the necessary condition for performing SemCom is that the two VUEs (transmitter and receiver) hold

[2] Other known probability distributions can also be adopted without changing the remaining modeling and solution.

[3] The KB popularity rank of each VUE can be estimated based on its historical messaging records, which will not be discussed in this paper.

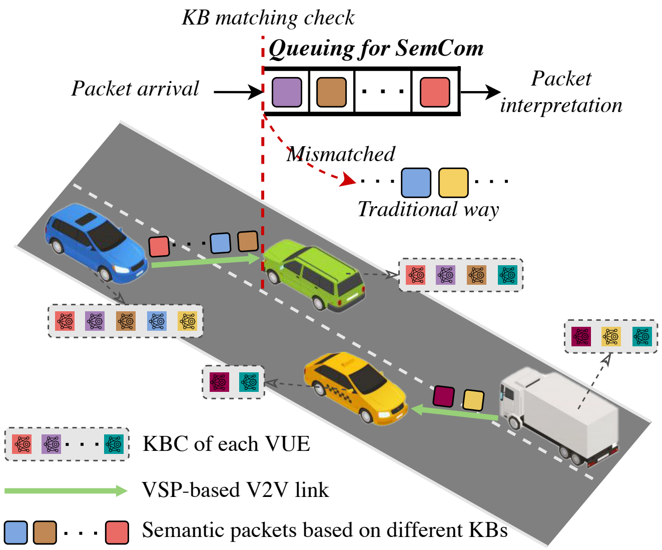

Fig. 2. The knowledge matching based queuing model for semantic data packets transmitted between VUEs in the SCVN.

common KBs. Moreover, the single association is required for all VUEs in the SCVN for practical purposes, i.e., each VUE can be paired with only one (another) VUE at a time.

Let $\beta_{i\looparrowright j}$ denote the binary VSP indicator for a VUE i-VUE j pair (suppose that VUE i is the sender and VUE j is the receiver), where

$$\beta_{i\looparrowright j}=\begin{cases}1, & \text{if VUE } i \text{ is associated with VUE } j;\\ 0, & \text{otherwise.}\end{cases} \tag{3}$$

Note that the presented communication performance (such as latency, reliability, and throughput) should be different when swapping the roles of sender and receiver in the same VUE pair, since the KBs utilized for SemCom are determined by the sender's preference. For this reason, we use the notation $\looparrowright$ here as an auxiliary illustration to specify the roles of the sender and receiver in each VUE pair.

D. Knowledge Matching Based Queuing Model

As depicted in Fig. 2, the knowledge matching based semantic packet queuing delay is employed as the latency metric of SemCom, to characterize the average sojourn time of semantic data packets in the receiver VUE's queue buffer (following the first-come first-serve rule). For better illustration, three major differences between SemCom-based and traditional communication-based queuing models are listed below: 1) Each semantic data packet is associated with a specific service type, i.e., a certain KB; 2) Semantic data packets generated based on different KBs can co-exist in the queue, and have independent average arrival rate and interpretation time; 3) Not all semantic data packets arriving at the receiver VUE are always allowed to enter its queue, as some of them may mismatch the KBs currently held, rendering these packets uninterpretable. To avoid pointless queuing, these mismatched packets may have to choose traditional communication channels for information transfer, and will not be counted in the arrival process of the queue.

To preserve generality, we first suppose a Poisson data arrival process with average rate $\lambda_i^n=\lambda_i p_i^n$ for a sender VUE i to account for its local semantic packet generation based on KB n, where λ_i is the total arrival rate of all semantic packets at VUE i. In line with this, we can obtain the overall arrival rate of semantic packets from sender VUE i to receiver VUE j as $\sum_{n\in\mathcal{K}}\alpha_i^n\lambda_i^n$, and the effective arrival rate of semantic packets (i.e., these KB-matched semantic packets) in the queue is given as $\sum_{n\in\mathcal{K}}\alpha_i^n\alpha_j^n\lambda_i^n$, thereby the arrival rate of mismatched semantic packets should be the value of the former minus the latter, that is, $\sum_{n\in\mathcal{K}}\alpha_i^n\left(1-\alpha_j^n\right)\lambda_i^n$. Herein, denoting the ratio of the arrival rate of mismatched semantic packets to the arrival rate of all received semantic packets at VUE i-VUE j pair as $\theta_{i\looparrowright j}$, namely *knowledge mismatch degree*, which is explicitly calculated by

$$\theta_{i\looparrowright j}=\frac{\sum_{n\in\mathcal{K}}\alpha_i^n\left(1-\alpha_j^n\right)\lambda_i^n}{\sum_{n\in\mathcal{K}}\alpha_i^n\lambda_i^n}. \tag{4}$$

In parallel, let a random variable I_j^n denote the Markovian interpretation time [18] required by KB n-based packets at VUE j with mean $1/\mu_j^n$, which is determined by the computing capability of the vehicle and the type of the desired KB. However, since multiple packets based on different KBs are allowed to queue at the same time, it is seen that the interpretation time distribution for a receiver VUE should be treated as a general distribution [19]. If further taking into account the KB popularity, we can calculate the ratio of the amount of KB n-based packets to the total packets in the VUE i-VUE j pair's queue by $\epsilon_{i\looparrowright j}^n=p_i^n/\sum_{f\in\mathcal{K}}\alpha_i^f\alpha_j^f p_i^f$. With the independence among packets based on different KBs, the interpretation time required by each packet in the queue is now expressed as $W_{i\looparrowright j}=\sum_{n\in\mathcal{K}}\alpha_i^n\alpha_j^n\epsilon_{i\looparrowright j}^n I_j^n$.

Since the Markovian arrival process leads to the correlated packet arrivals while the service pattern of packets obeys a general distribution, the queue of each VUE pair in the SCVN can be modeled as an M/G/1 system, which has been widely used to model data traffic in wireless networks. According to the *Pollaczek-Khintchine formula* [20], the average queuing latency for the VUE i-VUE j pair, denoted as $\delta_{i\looparrowright j}$, is determined as follows[4]

$$\delta_{i\looparrowright j}=\frac{\lambda_{i\looparrowright j}^{eff}\cdot\left(\mathbb{E}^2\left[W_{i\looparrowright j}\right]+Var\left(W_{i\looparrowright j}\right)\right)}{2\left(1-\lambda_{i\looparrowright j}^{eff}\cdot\mathbb{E}\left[W_{i\looparrowright j}\right]\right)}. \tag{5}$$

On this basis, again leveraging the independence of I_j^n over n, we can then obtain the expectation of the interpretation time for all semantic data packets in the queue by

$$\mathbb{E}\left[W_{i\looparrowright j}\right]=\sum_{n\in\mathcal{K}}\alpha_i^n\alpha_j^n\epsilon_{i\looparrowright j}^n\mathbb{E}\left[I_j^n\right]=\sum_{n\in\mathcal{K}}\frac{\alpha_i^n\alpha_j^n\epsilon_{i\looparrowright j}^n}{\mu_j^n}, \tag{6}$$

[4]In order to guarantee the steady-state of the queuing system, a condition of $\lambda_{i\looparrowright j}^{eff}\mathbb{E}\left[W_{i\looparrowright j}\right]<1$ must be satisfied before proceeding [19]. In this work, we assume that the packet interpretation rate is larger than the packet arrival rate to make the queuing latency finite and thus solvable.

and the variance of $W_{i \leftrightarrow j}$ is given by

$$\begin{aligned} Var\left(W_{i \leftrightarrow j}\right) &= \sum_{n \in \mathcal{K}} \alpha_i^n \alpha_j^n \left(\epsilon_{i \leftrightarrow j}^n\right)^2 Var\left[I_j^n\right] \\ &= \sum_{n \in \mathcal{K}} \alpha_i^n \alpha_j^n \left(\frac{\epsilon_{i \leftrightarrow j}^n}{\mu_j^n}\right)^2. \end{aligned} \tag{7}$$

Based on (6) and (7), $\delta_{i \leftrightarrow j}$ in (5) can be rewritten in (8), as shown at the bottom of this page.

III. Problem Formulation

In line with the xURLLC requirements, it is of paramount importance to achieve the optimality of the overall queuing delay in the SCVN, while being subject to several SemCom-relevant reliability requirements as well as practical system constraints. To that end, we identify and formulate a latency-minimization problem in a joint optimization manner of the KBC indicator α_i^n and the VSP indicator $\beta_{i \leftrightarrow j}$. For ease of illustration, we define a matrix $\boldsymbol{\alpha} = \{\alpha_i^n | i \in \mathcal{V}, n \in \mathcal{K}\}$ and a matrix $\boldsymbol{\beta} = \{\beta_{i \leftrightarrow j} | i \in \mathcal{V}, j \in \mathcal{V}_i\}$ consisting of all binary variables of KBC and VSP, respectively. On the basis of these, our joint optimization problem is formulated as follows:

$$\begin{aligned} \mathbf{P0}: \min_{\boldsymbol{\alpha}, \boldsymbol{\beta}} \quad & \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}_i} \beta_{i \leftrightarrow j} \delta_{i \leftrightarrow j} && (9) \\ \text{s.t.} \quad & \sum_{n \in \mathcal{K}} \alpha_i^n \cdot s_n \leqslant C_i, \ \forall i \in \mathcal{V}, && (9a) \\ & \eta_i \geqslant \eta_0, \ \forall i \in \mathcal{V}, && (9b) \\ & \sum_{j \in \mathcal{V}_i} \beta_{i \leftrightarrow j} = 1, \ \forall i \in \mathcal{V}, && (9c) \\ & \beta_{i \leftrightarrow j} = \beta_{j \leftrightarrow i}, \ \forall (i, j) \in \mathcal{V} \times \mathcal{V}_i, && (9d) \\ & \sum_{j \in \mathcal{V}_i} \beta_{i \leftrightarrow j} \theta_{i \leftrightarrow j} \leqslant \theta_0, \ \forall i \in \mathcal{V}, && (9e) \\ & \alpha_i^n \in \{0, 1\}, \ \forall (i, n) \in \mathcal{V} \times \mathcal{K}, && (9f) \\ & \beta_{i \leftrightarrow j} \in \{0, 1\}, \ \forall (i, j) \in \mathcal{V} \times \mathcal{V}_i. && (9g) \end{aligned}$$

Constraint (9a) ensures that the total size of KBs constructed at each vehicle cannot exceed its maximum storage capacity, while constraint (9b) corresponds to the aforementioned knowledge preference satisfaction requirement for each VUE. Constraints (9c) and (9d) mathematically model the single-association requirement of VUEs. Constraint (9e) represents that the knowledge mismatch degree of each VUE pair should not be over the threshold θ_0, which guarantees sufficiently high reliability of semantic information delivery. Constraints (9f) and (9g) characterize the binary properties of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, respectively.

Carefully examining $\mathbf{P0}$, it is seen that addressing this problem is rather challenging due to several inevitable mathematical obstacles. First of all, $\mathbf{P0}$ is an NP-hard optimization problem as demonstrated below. Consider a special case of $\mathbf{P0}$ where all $\boldsymbol{\beta}$-related constraints are satisfied. In this case, constraints (9c)-(9e) and (9g) can all be removed, and the primal problem degenerates into a classical 0-1 multi-knapsack problem that is known to be NP-hard [21]. Hence, $\mathbf{P0}$ is also NP-hard. Another nontrivial point originates from the complicated objective function, which prevents us from using the conventional two-step solution (i.e., relaxation and recovery) to approach optimality. In more detail, the problem after relaxing $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ should still be a nonconvex optimization problem owing to the non-convexity preserved in the objective function (9) and constraint (9e). Therefore, a severe performance penalty will be incurred from the procedure of integer recovery due to the huge performance compromise on solving the nonconvex problem for relaxed variables [22]–[24]. In view of the above mathematical challenges, we propose an efficient solution S^4 in the subsequent section to solve $\mathbf{P0}$ and obtain the joint optimal KBC and VSP solution.

IV. Proposed SemCom-Empowered Service Supplying Solution (S^4)

In this section, we illustrate how to design our proposed solution S^4 to cope with the SemCom-empowered service provisioning problem $\mathbf{P0}$ in vehicular networks. As depicted in Fig. 3, a Lagrange dual method is first leveraged to eliminate the cross-term constraints in $\mathbf{P0}$ with a corresponding dual optimization problem transformed (referring to $\mathbf{D0}$ in Section IV.A). Then given the dual variable in each iteration, we dedicatedly develop a two-stage method to determine $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ for the dual problem, where the optimality will be theoretically proved in Section IV.B. Specifically, in the first stage, we subtly construct U ($U = \left(\sum_{i \in \mathcal{V}} |\mathcal{V}_i|\right)/2$) subproblems (referring to $\mathbf{P1}_{i,j}, \forall (i,j) \in \mathcal{V} \times \mathcal{V}_i, j > i$), each of which aims to independently seek the optimal KBC sub-policy (with respect to only α_i^n and α_j^n) for each individual VUE pair (as detailed in Section IV.C). After solving all the U subproblems, the optimal coefficient matrix $\boldsymbol{\Omega}$ is obtained for all potential VUE pairs in the SCVN, by which we further construct a new subproblem (referring to $\mathbf{P2}$) in the second stage to find the optimal VSP strategy for $\boldsymbol{\beta}$ (as detailed in Section IV.D). In the end, we present the workflow of S^4 along with its complexity analysis in Section IV.E.

A. Primal-Dual Problem Transformation

We first incorporate constraint (9e) into the objective function (9) by associating a Lagrange multiplier $\boldsymbol{\tau} = \{\tau_i | i \in \mathcal{V}\}$.

$$\delta_{i \leftrightarrow j} = \frac{\left[\left(\sum_{n \in \mathcal{K}} \alpha_i^n \alpha_j^n \frac{p_i^n / \mu_j^n}{\sum_{f \in \mathcal{K}} \alpha_i^f \alpha_j^f p_i^f}\right)^2 + \sum_{n \in \mathcal{K}} \alpha_i^n \alpha_j^n \left(\frac{p_i^n / \mu_j^n}{\sum_{f \in \mathcal{K}} \alpha_i^f \alpha_j^f p_i^f}\right)^2\right] \cdot \left(\sum_{n \in \mathcal{K}} \alpha_i^n \alpha_j^n \lambda_i^n\right)}{2\left[1 - \left(\sum_{n \in \mathcal{K}} \alpha_i^n \alpha_j^n \lambda_i^n\right) \cdot \left(\sum_{n \in \mathcal{K}} \alpha_i^n \alpha_j^n \frac{p_i^n / \mu_j^n}{\sum_{f \in \mathcal{K}} \alpha_i^f \alpha_j^f p_i^f}\right)\right]}. \tag{8}$$

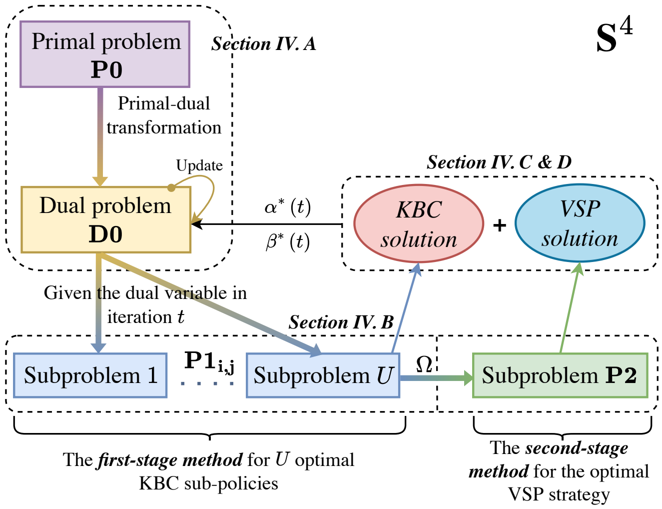

Fig. 3. Illustration of the proposed solution S^4.

That way, the associated Lagrange function is obtained by

$$\begin{aligned} &L\left(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\tau}\right) \\ &= \sum_{i\in\mathcal{V}}\sum_{j\in\mathcal{V}_i}\beta_{i\looparrowright j}\delta_{i\looparrowright j} + \sum_{i\in\mathcal{V}}\tau_i\left(\sum_{j\in\mathcal{V}_i}\beta_{i\looparrowright j}\theta_{i\looparrowright j} - \theta_0\right) \\ &= \sum_{i\in\mathcal{V}}\sum_{j\in\mathcal{V}_i}\beta_{i\looparrowright j}\left(\delta_{i\looparrowright j} + \tau_i\theta_{i\looparrowright j}\right) - \theta_0\sum_{i\in\mathcal{V}}\tau_i \\ &\triangleq \widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha}, \boldsymbol{\beta}) - \theta_0\sum_{i\in\mathcal{V}}\tau_i, \end{aligned} \tag{10}$$

where $\widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is defined for expression brevity. Then, the Lagrange dual problem of $\mathbf{P0}$ should be formulated as

$$\mathbf{D0}: \max_{\boldsymbol{\tau}} \quad D\left(\boldsymbol{\tau}\right) = g_{\boldsymbol{\alpha},\boldsymbol{\beta}}\left(\boldsymbol{\tau}\right) - \theta_0\sum_{i\in\mathcal{V}}\tau_i \tag{11}$$

$$\text{s.t.} \quad \tau_i \geqslant 0, \ \forall i \in \mathcal{V}, \tag{11a}$$

where we have

$$\begin{aligned} g_{\boldsymbol{\alpha},\boldsymbol{\beta}}\left(\boldsymbol{\tau}\right) \ &= \ \inf_{\boldsymbol{\alpha},\boldsymbol{\beta}} \ \widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha}, \boldsymbol{\beta}) \\ &\text{s.t.} \quad (9\text{a}) - (9\text{d}), (9\text{f}), (9\text{g}). \end{aligned} \tag{12}$$

Notably, the optimality of the convex problem $\mathbf{D0}$ gives at least the best lower bound of $\mathbf{P0}$, even if $\mathbf{P0}$ is nonconvex, according to the duality property [25]. Hence, our focus now naturally shifts to seeking the optimal solution to $\mathbf{D0}$.

Given the initial dual variable $\boldsymbol{\tau}$, we can solve problem (12) in the first place to find the optimal solution $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, the details of which will be presented in the next subsection. After that, a subgradient method is employed for updating $\boldsymbol{\tau}$ to solve $\mathbf{D0}$ in an iterative fashion, as shown in Fig. 3. Specifically, the partial derivatives with respect to $\boldsymbol{\tau}$ in the objective function $D\left(\boldsymbol{\tau}\right)$ are set as the subgradient direction in each iteration. Now suppose that in a certain iteration, say iteration t, each dual variable $\tau_i(t)$ $(i \in \mathcal{V})$ is updated as

$$\tau_i(t+1) = \left[\tau_i(t) - \nu(t)\cdot\left(\theta_0 - \sum_{j\in\mathcal{V}_i}\beta_{i\looparrowright j}\left(t\right)\theta_{i\looparrowright j}\left(t\right)\right)\right]^+. \tag{13}$$

The operator $[\cdot]^+$ here is to output the maximum value between its argument and zero, ensuring that $\boldsymbol{\tau}$ must be non-negative as constrained in (11a). $\nu(t)$ is the stepsize in iteration t and generally, the convergence of the subgradient descent method can be ensured with the proper stepsize [26].

B. Two-Stage Method Based on KBC and VSP

As discussed before, for a given $\boldsymbol{\tau}$ in each iteration, the optimal $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ need to be determined by solving problem (12). However, solving such a problem is still rather tricky due to the mathematical inseparability of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in the highly complex objective function $\widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha}, \boldsymbol{\beta})$. To this end, we propose a two-stage method to obtain the exactly optimal $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ with a low computational complexity.

In the first stage, we focus on multiple independent KB construction subproblems, each corresponding to a potential VUE pair in the SCVN. In particular, here the performances of the VUE i-VUE j pair (i.e., the sender VUE i and the receiver VUE j) and the VUE j-VUE i pair (i.e., the sender VUE j and the receiver VUE i) need to be considered together, and for ease of distinction, we refer to the two as a *VUE* i,j *pair*, $\forall\left(i,j\right) \in \mathcal{V}\times\mathcal{V}_i, j > i$. In other words, for any KBC subproblem, we have $\beta_{i\looparrowright j} = \beta_{j\looparrowright i} = 1$ in $\widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ corresponding to a given VUE i,j pair, while all other VUE pairs are not considered. Therefore, different KBC subproblems can be solved independently, and in this way, let

$$\omega_{i,j} = \left(\delta_{i\looparrowright j} + \tau_i\theta_{i\looparrowright j}\right) + \left(\delta_{j\looparrowright i} + \tau_j\theta_{j\looparrowright i}\right). \tag{14}$$

Obviously, we have $\omega_{i,j} = \omega_{j,i}$, thus only one case of $j > i$ needs to be investigated for each potential VUE i,j pair.

In this context, we now construct $U = \left(\sum_{i\in\mathcal{V}}|\mathcal{V}_i|\right)/2$ subproblems, each of which is denoted as $\mathbf{P1}_{i,j}$ to seek the optimal KBC sub-policy only for an individual VUE i,j pair. Herein, it is worth pointing out that the optimal KBC solution to problem (12) cannot be achieved by simply combining the obtained sub-policies of these $\mathbf{P1}_{i,j}$, but these sub-policies will be used to construct the subsequent VSP subproblem to finalize the joint optimal solution of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ for (12). Given the dual variable $\boldsymbol{\tau}$ in each iteration,[5] $\mathbf{P1}_{i,j}$ becomes

$$\mathbf{P1}_{i,j}: \min_{\{\alpha_i^n\},\{\alpha_j^n\}} \quad \omega_{i,j} \tag{15}$$

$$\text{s.t.} \quad \sum_{n\in\mathcal{K}}\alpha_i^n\cdot s_n \leqslant C_i, \tag{15a}$$

$$\sum_{n\in\mathcal{K}}\alpha_j^n\cdot s_n \leqslant C_j, \tag{15b}$$

$$\eta_i \geqslant \eta_0, \ \eta_j \geqslant \eta_0, \tag{15c}$$

$$\alpha_i^n \in \{0,1\}, \alpha_j^n \in \{0,1\}, \ \forall n \in \mathcal{K}. \tag{15d}$$

By solving $\mathbf{P1}_{i,j}$,[6] we can obtain the optimal KBC sub-policies for VUE i (denoted as $\boldsymbol{\alpha}^*_{i_{(j)}}$) and VUE j (denoted as $\boldsymbol{\alpha}^*_{j_{(i)}}$),[7] corresponding to the individual VUE i,j pair. The

[5]For simplicity, we omit $\boldsymbol{\tau}$ from all notations associated with $\mathbf{P1}_{i,j}$ and $\mathbf{P2}$ in this paper.

[6]The solution details of $\mathbf{P1}_{i,j}$ as well as $\mathbf{P2}$ will be introduced in the subsequent Subsection C and D, respectively.

[7]For auxiliary illustration, we use $(\cdot)$ in the subscript to specify the VUE pair attribute (relation) for each VUE's KBC sub-policy obtained from $\mathbf{P1}_{i,j}$.

following proposition explicitly shows how the sub-policy of $\mathbf{P1}_{i,j}$ correlates to the solution of problem (12).

Proposition 1. *Let $\boldsymbol{\alpha}^* = [\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*, \cdots, \boldsymbol{\alpha}_V^*]^T$ be the optimal KBC solution to the problem in (12) given the dual variable $\boldsymbol{\tau}$, where $\boldsymbol{\alpha}_i^*$ represents the optimal KBC policy of VUE i. Then we have $\forall i \in \mathcal{V}$, $\exists j \in \mathcal{V}_i$, such that $\boldsymbol{\alpha}_{i_{(j)}}^* = \boldsymbol{\alpha}_i^*$.*

Proof: Please see Appendix A. ■

From Proposition 1, it is observed that the optimal KBC policy of each VUE can be found by solving a certain $\mathbf{P1}_{i,j}$. Hence, considering the single-association requirement of V2V pairing, the optimal VSP strategy becomes the only key to finalize the optimal solution to (12). To achieve this, we first obtain the optimal coefficient matrix for $\boldsymbol{\beta}$ in (12) to account for all VSP possibilities. By calculating optimum $\omega_{i,j}$ (denoted as $\omega_{i,j}^*, \forall (i,j) \in \mathcal{V} \times \mathcal{V}_i$) in $\mathbf{P1}_{i,j}$, the optimal coefficient matrix is formed as

$$\boldsymbol{\Omega} = \begin{bmatrix} +\infty & \omega_{1,2}^* & \omega_{1,3}^* & \cdots & \omega_{1,V}^* \\ \omega_{2,1}^* & +\infty & \omega_{2,3}^* & \cdots & \omega_{2,V}^* \\ \omega_{3,1}^* & \omega_{3,2}^* & +\infty & \cdots & \omega_{3,V}^* \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \omega_{V,1}^* & \omega_{V,2}^* & \omega_{V,3}^* & \cdots & +\infty \end{bmatrix}. \tag{16}$$

$\boldsymbol{\Omega}$ is a $V \times V$ symmetric matrix where $\omega_{i,j}^* = \omega_{j,i}^*$, and all elements on its main diagonal are set to $+\infty$ to indicate the fact that a VUE cannot communicate with itself, i.e., $j \neq i$. Besides, note that some $\omega_{i,j}^*$s in $\boldsymbol{\Omega}$ also have a value $+\infty$ if VUE j is not the direct neighbor of VUE i, i.e., $j \notin \mathcal{V}_i$.

Next, we concentrate upon the optimal vehicle service pairing strategy by constructing a new subproblem in the second stage. In line with the objective $\widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ and $\boldsymbol{\beta}$-related constraints in (12), the VSP subproblem is written as

$$\mathbf{P2}: \min_{\boldsymbol{\beta}} \quad \frac{1}{2} \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}_i} \beta_{i \looparrowright j} \omega_{i,j}^* \tag{17}$$

$$\text{s.t.} \quad (9c), (9d), (9g). \tag{17a}$$

Given any $\boldsymbol{\tau}$, the optimal VSP strategy $\boldsymbol{\beta}$ (denoted as $\boldsymbol{\beta}^* = \left[\beta_{1 \looparrowright j_1^*}, \beta_{2 \looparrowright j_2^*}, \cdots, \beta_{V \looparrowright j_V^*}\right]^T$) can be directly finalized by solving $\mathbf{P2}$, where $\beta_{i \looparrowright j_i^*}$ ($\forall i \in \mathcal{V}$) indicates that VUE j_i^* is the optimal SemCom node for VUE i, i.e., $\beta_{i \looparrowright j_i^*} = 1$. Afterward, we feed back the obtained $\boldsymbol{\beta}^*$ to $\boldsymbol{\Omega}$ to further finalize the optimal KBC policy $\boldsymbol{\alpha}^*$ for all VUEs. In the context of the solution to $\mathbf{P1}_{i,j}$, the approach to finalize $\boldsymbol{\alpha}^*$ can be stated more precisely as: for any $i \in \mathcal{V}$, we have $\boldsymbol{\alpha}_i^* = \boldsymbol{\alpha}_{i_{(j_i^*)}}^*$. The rationale behind this is established in accordance with the following proposition.

Proposition 2. *Given any dual variable $\boldsymbol{\tau}$, $(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*)$ is exactly the optimal solution to the problem in (12).*

Proof: Please see Appendix B. ■

From Proposition 2, it is seen that for problem (12) in each iteration, the proposed two-stage method is ensured to find the optimal solution. Apart from this, the optimization difficulty of each subproblem, either $\mathbf{P1}_{i,j}$ or $\mathbf{P2}$, is considered to be greatly decreased due to the reduced number of the optimization variables. In what follows, we will present our optimal solutions to $\mathbf{P1}_{i,j}$ and $\mathbf{P2}$, respectively.

C. Near-Optimal Solution to $\mathbf{P1}_{i,j}$

Carefully examining $\mathbf{P1}_{i,j}, \forall (i,j) \in \mathcal{V} \times \mathcal{V}_i, j > i$, it can be observed that $\delta_{i \looparrowright j}$ mainly makes the objective function $\omega_{i,j}$ still highly complex and nonconvex with two binary variables $\boldsymbol{\alpha}_i$ and $\boldsymbol{\alpha}_j$. To that end, a modified metaheuristic algorithm based on tabu search (TS) is employed here to efficiently determine a near-optimal KB construction policy for two VUEs in the VUE i,j pair with the consideration of SemCom features. In detail, the KBC solution is illustrated as follows:

- *Initial Feasible Solution Generation:* As the iterative search algorithm, an initial feasible solution (denoted as a $2N$-dimensional vector $\boldsymbol{\alpha}_{i,j}^I$) is needed as the search starting point [27]. To speed up convergence and enhance optimization performance, we heuristically adopt a KB preference and KB matching-aware approach to generate $\boldsymbol{\alpha}_{i,j}^I$ for a better initial solution performance. More concretely, first let two N-dimensional vectors $\boldsymbol{\alpha}_i^I$ and $\boldsymbol{\alpha}_j^I$ denote the KBC solutions of VUE i and VUE j, respectively, with all elements being 0 for initialization. Meanwhile, suppose there are two variable sets, denoted as $\hat{\mathcal{K}}$ and $\check{\mathcal{K}}$, to record KB-relevant information, where $\hat{\mathcal{K}} = \mathcal{K}$ and $\check{\mathcal{K}} = \emptyset$ are initialized. With these, we attempt to find a KB n_0 with the highest sum of KB preferences of the two VUEs by

$$n_0 = \arg\max_{n \in \hat{\mathcal{K}}} \left(p_i^n + p_j^n\right). \tag{18}$$

Then, in order to meet the knowledge mismatch requirement, the values of both $\boldsymbol{\alpha}_i^I$ and $\boldsymbol{\alpha}_j^I$ are updated as

$$\alpha_i^{n_0} = 1 \quad \text{and} \quad \alpha_j^{n_0} = 1. \tag{19}$$

Next, we let $\hat{\mathcal{K}} = \hat{\mathcal{K}} \backslash n_0$ and $\check{\mathcal{K}} = \check{\mathcal{K}} \cup \{n_0\}$, and then repeat the two procedures in (18) and (19) until both VUEs satisfy the minimum knowledge preference satisfaction requirement in constraint (15c). However, notice that constraint (15a) or (15b) may be violated during the above process, in which case we need to find the maximum-size KB in $\check{\mathcal{K}}$ by

$$n_1 = \arg\max_{n \in \check{\mathcal{K}}} s_n, \tag{20}$$

and then reset the corresponding KBC indicators in $\boldsymbol{\alpha}_i^I$ and $\boldsymbol{\alpha}_j^I$ to 0, i.e.,

$$\alpha_i^{n_1} = 0 \quad \text{and} \quad \alpha_j^{n_1} = 0. \tag{21}$$

As a result, we obtain an initial feasible solution

$$\boldsymbol{\alpha}_{i,j}^I = \left[\boldsymbol{\alpha}_i^I, \boldsymbol{\alpha}_j^I\right]. \tag{22}$$

- *Neighboring Solution Searching:* Let a $2N$-dimensional vector $\boldsymbol{\alpha}_{i,j}^C$ store the current solution in each search iteration, and $\mathcal{H}\left(\boldsymbol{\alpha}_{i,j}^C\right)$ denote its neighboring solution set, which should not include solutions that are already recorded in a tabu list, denoted as a set $\mathcal{I}$ (which will be explained later). Naturally, our $\mathcal{H}\left(\boldsymbol{\alpha}_{i,j}^C\right)$ is defined as

$$\mathcal{H}(\boldsymbol{\alpha}_{i,j}^C) = \{\boldsymbol{\alpha}_{i,j} \colon \|\boldsymbol{\alpha}_{i,j} - \boldsymbol{\alpha}_{i,j}^C\| \leqslant \sigma, \\ \boldsymbol{\alpha}_{i,j} \notin \mathcal{I}, \boldsymbol{\alpha}_{i,j} \in \psi\}, \tag{23}$$

where σ is the size of the maximum neighborhood space, and ψ represents the feasible solution space of $\mathbf{P1}_{i,j}$. Based on the above definitions, we are able to find the best solution within the current $\mathcal{H}\left(\boldsymbol{\alpha}_{i,j}^{C}\right)$ that can yield the minimum value of $\omega_{i,j}$ in an iterative fashion.

- *Tabu List Update:* The tabu list $\mathcal{I}$ is a special memory mechanism for preventing subsequent searches from looping back to previously visited solutions so as to avoid trapping into the local optimum [28]. In this way, whenever there is a newly obtained $\boldsymbol{\alpha}_{i,j}^{C}$, it should be added into $\mathcal{I}$ (cannot exceed its given maximum length [27]). Besides, let $\boldsymbol{\alpha}_{i,j}^{*}$ denote another vector dedicated to storing the best solution obtained so far. Particularly, once $\boldsymbol{\alpha}_{i,j}^{C}$ is found to be better than $\boldsymbol{\alpha}_{i,j}^{*}$ in any iteration, it will not be added into $\mathcal{I}$ but we will have

$$\boldsymbol{\alpha}_{i,j}^{*}=\boldsymbol{\alpha}_{i,j}^{C}. \tag{24}$$

- *Algorithm Termination Check:* Before commencing a new iteration, an algorithm termination criterion needs to be checked, which can be either a maximum iteration restriction or a performance improvement threshold of $\boldsymbol{\alpha}_{i,j}^{*}$ under a certain number of consecutive iterations.

D. Optimal Solution to $\mathbf{P2}$

After the KBC sub-policy $\boldsymbol{\alpha}_{i,j}^{*}$ is found for each potential VUE i,j pair, we can determine the optimal coefficient matrix $\boldsymbol{\Omega}$ to solve the VSP subproblem $\mathbf{P2}$. Since its objective function and two equality constraints (9c) and (9d) are all linear, the only challenge is the 0-1 constraint in (9g).

With regards to this, we first relax $\boldsymbol{\beta}$ into a continuous variable between 0 and 1 to make $\mathbf{P2}$ a linear programming problem, which can be efficiently solved with toolboxes such as CVXPY [29]. Then the obtained continuous solution, denoted as $\boldsymbol{\beta}^{R}$, needs to be restored to the binary state under the original constraints. Here, we heuristically finalize the optimal VSP strategy $\boldsymbol{\beta}^{*}$ by

$$\beta_{i^{'}\looparrowright j^{'}}^{*}=\beta_{j^{'}\looparrowright i^{'}}^{*}=1, \tag{25}$$

if and only if

$$\left(i^{'},j^{'}\right)=\arg\max_{i\in\mathcal{V},j\in\mathcal{V}_i,j>i}\beta_{i\looparrowright j}^{R}. \tag{26}$$

Meanwhile, for the remaining $\beta_{i\looparrowright j}^{*}$ with respect to VUE $i^{'}$ and VUE $j^{'}$, we naturally have

$$\begin{cases}\beta_{i^{'}\looparrowright j}^{*}=\beta_{j\looparrowright i^{'}}^{*}=0, & \forall j\in\mathcal{V}_{i^{'}},j\neq j^{'}\\ \beta_{j^{'}\looparrowright i}^{*}=\beta_{i\looparrowright j^{'}}^{*}=0, & \forall i\in\mathcal{V}_{j^{'}},i\neq i^{'}\end{cases}. \tag{27}$$

Then we let $\mathcal{V}=\mathcal{V}\backslash\{i,j\}$, and repeat the above progresses until determining the optimal VSP solution for all VUEs. It can be seen that the number of variables is actually only $\left(\sum_{i\in\mathcal{V}}|\mathcal{V}_i|\right)/2$ when solving $\mathbf{P2}$, which is a fairly acceptable problem scale in practice. Hence, the performance compromise of our proposed heuristic VSP solution is believed to be small.

E. Workflow of S[4] and Complexity Analysis

In order to further demonstrate the full picture of the proposed solution S[4], we summarize the relevant technical points and present them in the following Algorithm 1.

Algorithm 1 The Proposed Solution S[4]

Input: The SCVN parameters s_n, C_i, r_i^n, ξ_i, λ_i, μ_i^n, η_0, θ_0
Output: The optimal KBC policy $\boldsymbol{\alpha}^{*}$ and the optimal VSP strategy $\boldsymbol{\beta}^{*}$ for each VUE $i,\forall i\in\mathcal{V}$

Initialize $t=0$, $\tau_i(0)$ and $\nu(0)$ to proper positive values;
Set the maximum number of iterations M for $\mathbf{D0}$;
while $t<M$ **do**
  for $i=1$ to V **do**
    for each $j\in\mathcal{V}_i$ **do**
      if $j>i$ **then**
        Determine the initial solution $\boldsymbol{\alpha}_{i,j}^{I}$ by (18)-(22);
        Initialize the TS iteration as $\tilde{t}=0$, the Tabu list $\mathcal{I}(0)=\emptyset$, and $\boldsymbol{\alpha}_{i,j}^{C}(0)=\boldsymbol{\alpha}_{i,j}^{*}(0)=\boldsymbol{\alpha}_{i,j}^{I}$;
        Set the neighborhood size σ and the maximum number of iterations Z for solving each $\mathbf{P1}_{i,j}$;
        while $\tilde{t}<Z$ **do**
          Determine $\mathcal{H}\left(\boldsymbol{\alpha}_{i,j}^{C}\left(\tilde{t}\right)\right)$ by (23);
          Find the best feasible solution in $\mathcal{H}\left(\boldsymbol{\alpha}_{i,j}^{C}\left(\tilde{t}\right)\right)$ and assign it to $\boldsymbol{\alpha}_{i,j}^{C}(\tilde{t}+1)$;
          if $\boldsymbol{\alpha}_{i,j}^{C}(\tilde{t}+1)$ is better than $\boldsymbol{\alpha}_{i,j}^{*}(\tilde{t})$ **then**
            Update $\boldsymbol{\alpha}_{i,j}^{*}(\tilde{t}+1)=\boldsymbol{\alpha}_{i,j}^{C}(\tilde{t}+1)$;
            Keep $\mathcal{I}\left(\tilde{t}+1\right)=\mathcal{I}\left(\tilde{t}\right)$;
          else
            Keep $\boldsymbol{\alpha}_{i,j}^{*}(\tilde{t}+1)=\boldsymbol{\alpha}_{i,j}^{*}(\tilde{t})$;
            Update $\mathcal{I}\left(\tilde{t}+1\right)=\mathcal{I}\left(\tilde{t}\right)\cup\left\{\boldsymbol{\alpha}_{i,j}^{C}(\tilde{t}+1)\right\}$;
          end if
          Update $\tilde{t}=\tilde{t}+1$;
        end while
        Calculate $\omega_{i,j}^{*}$ by substituting $\boldsymbol{\alpha}_{i,j}^{*}(Z)$ into (14);
        Assign $\omega_{j,i}^{*}=\omega_{i,j}^{*}$;
      end if
    end for
  end for
  Renew the optimal coefficient matrix $\boldsymbol{\Omega}\left(t\right)$ by (16);
  Solve the relaxed $\mathbf{P2}$ by CVXPY and obtain $\boldsymbol{\beta}^{R}\left(t\right)$;
  Finalize $\boldsymbol{\beta}^{*}\left(t\right)$ by (25)-(27);
  Finalize $\boldsymbol{\alpha}^{*}\left(t\right)$ by feeding $\boldsymbol{\beta}^{*}\left(t\right)$ back into $\boldsymbol{\Omega}\left(t\right)$;
  Update $\tau_i(t+1)$ by (13);
  Update $\nu(t+1)$ under a given rule;
  Update $t=t+1$;
end while

In line with this algorithm, the main flow of S[4] working in the practical SCVN is demonstrated as follows:

- *Network Initialization:* In the initial phase, each VUE i ($\forall i\in\mathcal{V}$) generates a Lagrange multiplier parameter τ_i and records all KBC-related status information, including its available KB storage (i.e., C_i), preferences for different KBs (i.e., $r_i^n,\forall n\in\mathcal{K}$, and ξ_i), and average local arrival rate as well as interpretation time for semantic data packets (i.e., λ_i and $1/\mu_i^n$). Then, all VUEs need to upload the above parameters to the RSU for subsequent implementation.

- *Optimal Policy Determination for KBC and VSP:* In this phase, the RSU first measures the SINR between VUEs to determine each VUE i's communication neighbors, i.e., $\mathcal{V}_i$. Having these, the RSU is capable of computing the current optimal KBC sub-policy for each individual VUE i, j pair ($j \in \mathcal{V}_i$) (referring to steps 4-29 in Algorithm 1). Afterward, the current optimal VSP and KBC policies can be jointly obtained in steps 30-33. Nevertheless, the Lagrange multiplier of each VUE is required to be updated with a given rule (steps 34-36), thus the RSU should repeat the procedures in steps 4-33 until satisfying the algorithm termination criterion, so as to finalize the optimal KBC and VSP policies $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}^*$.
- *SemCom-Empowered Service Provisioning:* Once each VUE receives the feedback information, it can request and download KBs from the RSU (according to $\boldsymbol{\alpha}^*$), and then pair with one of its neighbors (according to $\boldsymbol{\beta}^*$) to provide corresponding services for each other.

Herein, it is worth pointing out that the above workflow of S^4 is executed in a periodic timeline, and all decisions to update relevant parameters should be made at the end of each period. Besides, it can be observed that there are only four rounds of signaling interactions to implement a completed and successive KBC and VSP process, including vehicular information collection, optimal KBC and VSP policy assignment, KB downloading, and vehicle pairing. For each signaling interplay, only a few bits are needed to complete the functional confirmation work, hence the overall signaling overhead should be an apparently tolerable level in practice.

In terms of the computational complexity of S^4, it is first seen that for a single $\mathbf{P1}_{i,j}$, each feasible KBC solution within $\mathcal{H}\left(\boldsymbol{\alpha}_{i,j}^{C}\right)$ needs to be computed once in any of its iterations. Combining that σ is given as a small parameter compared to N in (23), the complexity in each iteration can be estimated by $\mathcal{O}\left(\binom{2N}{\sigma}\right) = \mathcal{O}\left(N^{\sigma}\right)$. If the maximum number of its iterations is assumed to be Z, then solving each $\mathbf{P1}_{i,j}$ would require complexity $\mathcal{O}\left(ZN^{\sigma}\right)$. In addition, the linear programming method is utilized for the relaxed $\mathbf{P2}$, where $\mathcal{O}\left(V^4\right)$ complexity is needed [30] to solve a group of $\left(\sum_{i\in\mathcal{V}}|\mathcal{V}_i|\right)$ VSP variables. Moreover, note that in any iteration of $\mathbf{D0}$, a total of $\left(\left(\sum_{i\in\mathcal{V}}|\mathcal{V}_i|\right)/2\right)$ subproblems $\mathbf{P1}_{i,j}$ and one subproblem $\mathbf{P2}$ need to be solved simultaneously, thereby its corresponding complexity is $\mathcal{O}\left(V^2ZN^{\sigma}+V^4\right)$. If denoting the maximum number of iterations that can make $\mathbf{D0}$ to converge as M, the proposed S^4 would have a polynomial-time overall complexity, given as $\mathcal{O}\left(MV^2\left(ZN^{\sigma}+V^2\right)\right)$.

V. Numerical Results and Discussions

In this section, numerical evaluations are conducted to demonstrate the performance of the proposed solution S^4 in SCVNs, where we employ Python 3.7-based PyCharm as the simulator platform and implement it in a computer with six CPU cores and Inter Core i7 processor. To preserve generality, we model a multi-lane freeway passing through a single cell with the RSU at its center, and multiple VUEs are dropped on the lanes according to the spatial Poisson process [32]. The main parameters of the basic network setup as well as their corresponding values can be found in Table I [31]–[35]. As for the settings relevant to SemCom, a total of 12 different KBs are preset to provide VUEs with a variety of distinct services, and each of them has a storage size randomly distributed from 1 to 5 units. Correspondingly, we set a uniform KB storage capacity of 24 units for all VUEs. Besides, each VUE's preference ranking for all KBs (i.e., r_i^n) is generated in an independent and random manner, where their respective Zipf distributions are assumed to have the same skewness 1.0. Likewise, either the average arrival rate of total semantic data packets, or the average interpretation time for packets based on the same KB n, is considered to be the same for all VUEs. Here, we fix the average total arrival rate λ_i at 100 packets/s and randomly generate the value of $1/\mu_i^n$ in a range of $5\times10^{-3} \sim 1\times10^{-2}$ s/packet with respect to different KB n, as the average interpretation time of different KBs-based packets is different from each other, which is to guarantee the steady-state of the queuing system at each VUE pair, as has mentioned in Footnote 4. Further, the minimum knowledge preference satisfaction threshold η_0 and the maximum knowledge mismatch degree threshold θ_0

TABLE I
Simulation Parameters

Parameters	**Values**
Number of VUEs (V)	60
Number of KBs (N)	12
Size of KB n (s_n)	$1 \sim 5$ units (randomly)
KB storage capacity of VUEs (C_i)	24 units
Skewness of the Zipf distribution with respect to each VUE's KB preference (ξ_i)	1.0
Average arrival rate of total semantic data packets of VUEs (λ_i)	100 packets/s [31]
Average interpretation time of KB n-based semantic data packets of VUEs ($1/\mu_i^n$)	$5\times10^{-3} \sim 1\times10^{-2}$ s/packet (randomly)
Cell radius of the RSU	500 m
VUE drop model	Spatial Poisson process [32]
Number of lanes	3 in each direction (6 in total)
Lane width	4 m
Absolute velocity of VUEs	70 km/h [33]
Density of VUEs	Average inter-vehicle distance is 2.5 sec $\times$ absolute velocity of VUEs [32]
Transmit power of VUEs	20 dBm [34]
Noise power	-114 dBm
Path loss model	$128.1 + 37.6\log\left(d\text{ [km]}\right)$ [35]
Channel fading model	Log-normal shadowing distribution with standard deviation of 8 dB + Rayleigh fast fading [35]
Minimum knowledge preference satisfaction threshold (η_0)	0.5
Maximum knowledge mismatch degree threshold (θ_0)	0.1

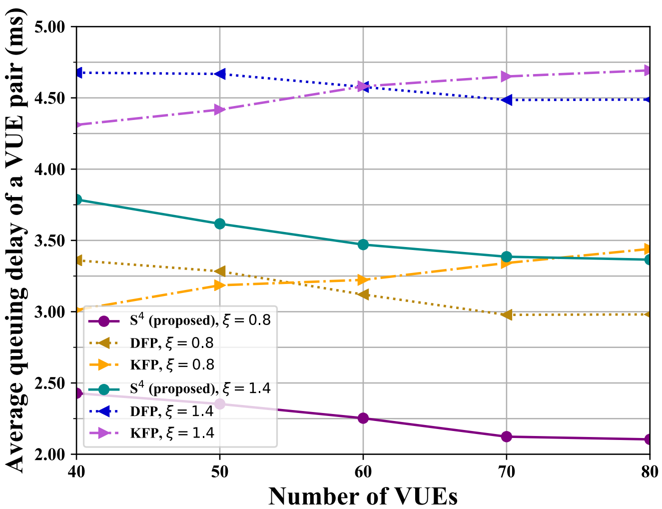

Fig. 4. Average queuing latency of a VUE pair vs. varying numbers of VUEs.

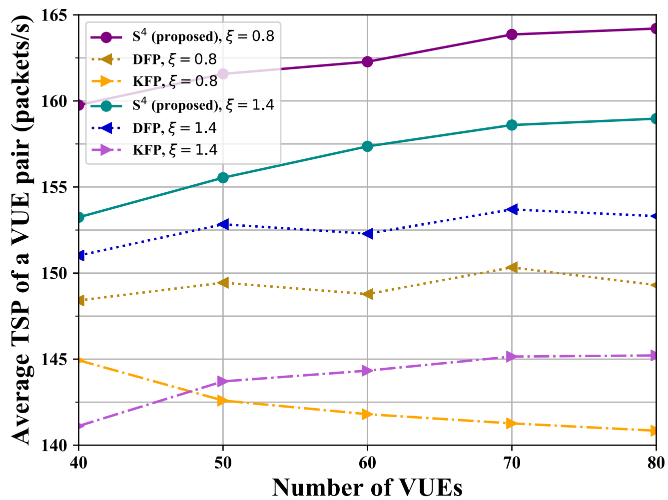

Fig. 5. Average TSP of a VUE pair vs. varying numbers of VUEs.

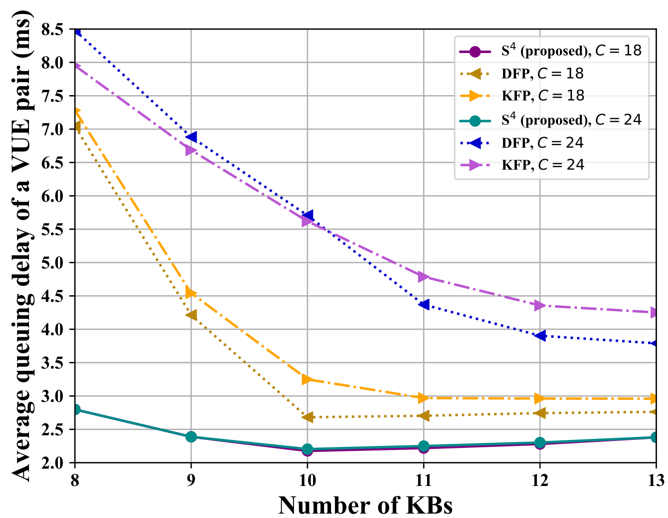

Fig. 6. Average queuing latency of a VUE pair with varying numbers of KBs.

are prescribed as 0.5 and 0.1, respectively. Notably, all these parameter values in Table I are set by default unless otherwise specified, and all subsequent numerical results are obtained by averaging over a sufficiently large number of trials.

For comparison purposes, we utilize two different benchmark schemes of SemCom-empowered service provisioning herein: 1) Distance-first pairing (DFP) strategy which assumes each VUE to choose its nearest unpaired VUE for V2V pairing; 2) Knowledge-first pairing (KFP) in which each VUE selects its neighboring unpaired VUE with the highest KB matching degree for V2V pairing. In the meantime, a personal preference-first KBC policy is considered for both benchmarks, which allows each VUE to construct KBs with the highest preferences until η_0 is satisfied, and then randomly select these unconstructed KBs until reaching respective maximum capacity.

Fig. 4 first depicts the average queuing latency performance of a VUE pair against varying numbers of VUEs, where two different KB preference skewness $\xi = 0.8$ and $\xi = 1.4$ are considered. In this figure, the latency of S^4 declines at the beginning with the number of VUEs, then remains stable beyond 70 VUEs, and it can always outperform both benchmarks with an average latency reduction of around 1 ms at any ξ. The rationale behind this trend is that the more neighbors each VUE can have, the better chance of achieving the low queuing latency for each VUE pair, which will be eventually stabilized when reaching the respective best achievable latency with a fixed bandwidth budget. Moreover, it is observed that a larger ξ causes a higher latency penalty, since the vast majority of VUEs' KBC is concentrated on a small number of KBs when ξ increases. Clearly, a larger ξ will make each participant more difficult to find the best VUE with the low latency under the given knowledge mismatching requirement θ_0, thus resulting in a degraded performance.

The above analysis also applies to Fig. 5, which compares all the three methods under the same settings as Fig. 4 to demonstrate the performance of the average throughput in semantic packets (TSP). Specifically, the TSP represents the total number of semantic packets that can be interpreted by a VUE pair per second, whose value is determined based on $\mathbb{E}\left[W_{i \leftrightarrow j}\right]$ in (6). Likewise, a higher TSP is obtained as the number of VUEs increases, and our S^4 is still far better than the two benchmarks at any point, e.g., with an average performance gain of 14 packets/s compared with DFP and 20 packets/s with KFP at $\xi = 0.8$. Again, we see a better TSP when the KB popularity is diluted by a smaller ξ.

Next, we explore the impact of varying number of KBs on the average queuing latency of a VUE pair with different VUE capacities $C = 18$ and $C = 24$, as demonstrated in Fig. 6. It can be found that the latency drops fast at the beginning, and then rises slightly after exceeding 10 KBs, whereas the performance of our S^4 still surpasses the benchmarks. This trend is attributed to the fact that more KBs imply less discrepancy in VUEs' preferences for different KBs given the fixed ξ, thereby at first leading to the higher probability for two paired VUEs constructing the KBs with high interpretation rates so as to render a lower delay. However, such performance

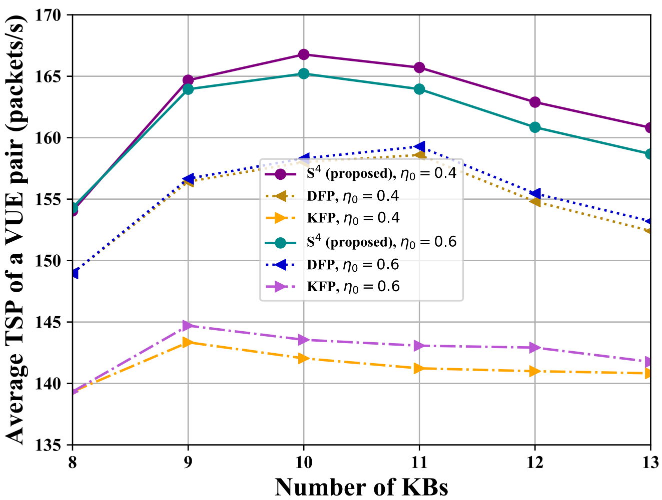

Fig. 7. Average TSP of a VUE pair with varying numbers of KBs.

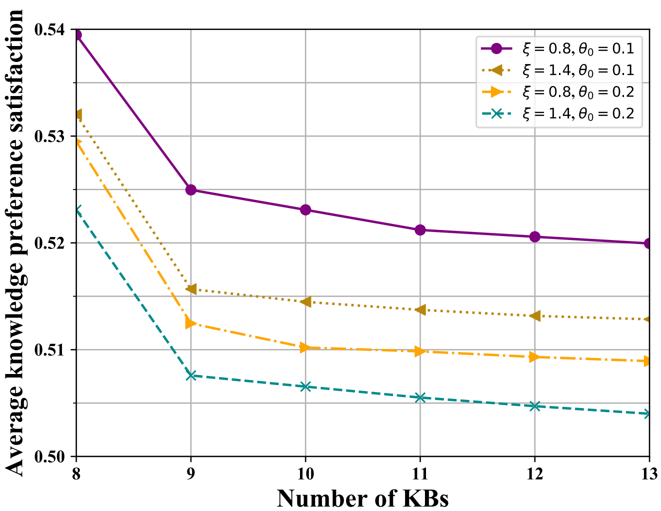

Fig. 8. Average knowledge preference satisfaction reached at a VUE with varying numbers of KBs.

gains will be saturated and even worsen when these KBs with low interpretation rates become inevitably dominant in order to meet the minimum knowledge preference satisfaction threshold η_0. Besides, it is seen that different VUE capacities have little effect on the latency of S^4, although the larger capacity can construct more KBs. This is due to the latency-minimization objective we particularly focus on in the delay-sensitive SCVN, and only the KBs with low interpretation time should be selected. Afterward, we draw the TSP performance against different numbers of KBs with $\eta_0 = 0.4$ and $\eta_0 = 0.6$, as shown in Fig. 7. The similar trend to Fig. 6 is observed here as well, i.e., the TSP of S^4 rises at the beginning and then falls after 10 KBs, and a much higher TSP is provided compared with the benchmarks. Meanwhile, it is noticed that a lower η_0 brings a better TSP, as fewer KBs need to be constructed to guarantee the high average interpretation rates in the queue, as mentioned earlier.

In addition, we validate the average knowledge preference satisfaction $\bar{\eta} = \frac{1}{V}\sum_{i\in\mathcal{V}} \eta_i$ reached at each VUE with varying

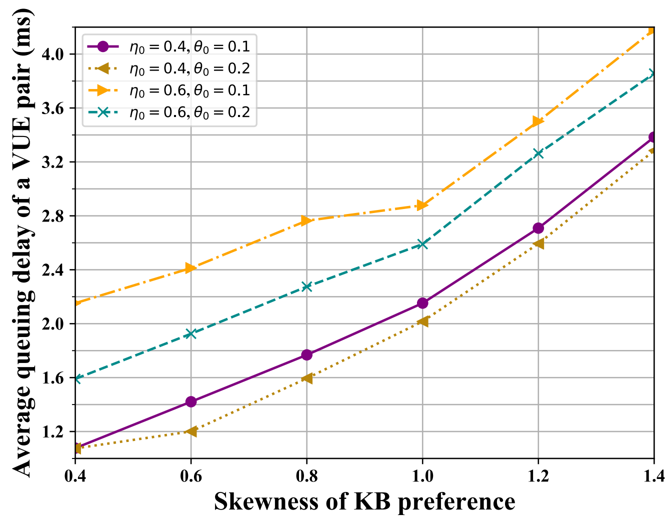

Fig. 9. Average queuing latency of a VUE pair with varying skewness of VUEs' KB preferences.

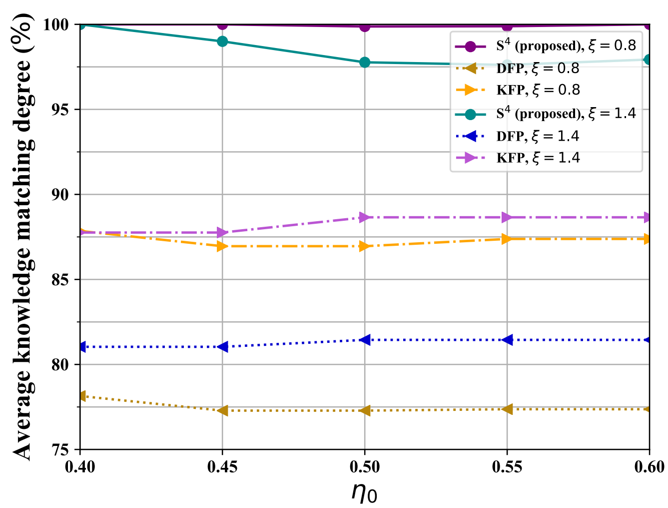

Fig. 10. Average knowledge matching degree of a VUE pair vs. different knowledge preference satisfaction requirements.

numbers of KBs as shown in Fig. 8, where $\xi = 0.8$, $\xi = 1.4$, $\theta_0 = 0.1$, and $\theta_0 = 0.2$ are taken into account. As the number of KBs increases, a lower $\bar{\eta}$ is obtained, which is to prevent these unnecessary KBs from being constructed while satisfying η_0 to the greatest extent. For the two curves with different ξ, referring to the analysis of Fig. 4, a higher ξ indicates a more concentrated KB preference, which means some extra KBs need to be constructed to meet the maximum θ_0 requirement. Because of this, we also see a lower $\bar{\eta}$ at a higher θ_0, since a more tolerable knowledge mismatch degree is more likely to avoid the unnecessary KBC.

Fig. 9 presents the effect of varying ξ on the average queuing delay compared between different η_0 and θ_0. As expected, the latency of all three methods increases with ξ, and such an upward trend is consistent with the previous results. Specially, it is observed that the curve of $\eta_0 = 0.4$ and $\theta_0 = 0.2$ brings the best latency performance. This can be understood as that either the lower satisfaction threshold or the higher mismatch tolerance can avoid the construction of unnecessary KBs with low interpretation rates, as discussed before. Naturally, the better latency performance is obtained when the constraints become less stringent.

Finally, we plot the average knowledge matching degree, defined as $\bar{\rho} = \frac{1}{V}\sum_{i\in\mathcal{V},j\in\mathcal{V}_i}(1-\theta_{i\looparrowright j})$, with two different ξ in Fig. 10. It can be seen that under the threshold of $\theta_0 = 0.1$, the proposed solution S^4 is always above a 97.5% match degree, which is much higher than that of benchmarks. Furthermore, a higher ξ causes the slight drop of $\bar{\rho}$, which exactly proves the conclusion of Fig. 4, i.e., a more concentrated KB preference makes it more difficult for VUEs to find a highly matching neighbor in the VSP phase, especially for the one that can bring lower latency and meet all constraints at the same time.

VI. Conclusions

In this paper, we proposed a novel solution S^4 to address the SemCom-empowered service provisioning problem in the SCVN. To align with the stringent xURLLC requirements, the KB matching based queuing latency expression of semantic data packets was first derived, and then we identified and formulated the fundamental problem of KBC and VSP to minimize the queuing latency for all VUE pairs. After the primal-dual problem transformation, a two-stage method was developed specifically to jointly solve multiple subproblems related to KBC and VSP with low computational complexity, and the solution optimality has been theoretically proved. Numerical results verified the sufficient performance superiority of S^4 in terms of both latency and reliability by comparing it with two different benchmarks.

This work can be served as a pioneer in exploring the potential of applying SemCom in vehicular networks for provisioning pertinent communication services with the aim of meeting the xURLLC requirements. Besides, other advanced networking issues in the SCVN, such as semantic-aware resource allocation or semantic transceiver design, can treat this work as the fundamental theoretical framework for reference. Since this work is limited to determining optimal instantaneous KBC and VSP policies for VUEs with known KB popularity, the further problem in an expanded SCVN scenario of considering high user mobility and unaware user preferences will be investigated in our future research.

Appendix A
Proof of Proposition 1

Given the optimal KBC solution $\boldsymbol{\alpha}^*$, let $\boldsymbol{\beta}^* = \left[\beta_{1\looparrowright j_1^*}, \beta_{2\looparrowright j_2^*}, \cdots, \beta_{V\looparrowright j_V^*}\right]^T$ be the corresponding optimal VSP solution to the problem in (12) under the same dual variable $\boldsymbol{\tau}$, where $\beta_{i\looparrowright j_i^*}$ ($\forall i \in \mathcal{V}$) indicates that VUE j_i^* is the optimal SemCom node for VUE i, i.e., $\beta_{i\looparrowright j_i^*} = 1$.

From $\omega_{i,j}$ defined in (14), the objective function $\widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha},\boldsymbol{\beta})$ in (12) can be rewritten as

$$\widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha},\boldsymbol{\beta}) = \frac{1}{2}\sum_{i\in\mathcal{V}}\sum_{j\in\mathcal{V}_i}\beta_{i\looparrowright j}\omega_{i,j} = \sum_{i\in\mathcal{V}}\sum_{j\in\mathcal{V}_i,j>i}\beta_{i\looparrowright j}\omega_{i,j}, \tag{28}$$

then we substitute $\boldsymbol{\beta}^*$ into (28) and yield

$$\widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha},\boldsymbol{\beta}^*) = \frac{1}{2}\sum_{i\in\mathcal{V}}\omega_{i,j_i^*} = \sum_{i\in\mathcal{V},i<j_i^*}\omega_{i,j_i^*}, \tag{29}$$

where ω_{i,j_i^*} is the term only related to VUE i, j_i^* pair.

Undoubtedly, if $\boldsymbol{\alpha}^*$ is further substituted into (29), we can straightforwardly reach the optimality of the problem in (12). Since different VUE i, j_i^* pairs are independent of each other, it means that different terms related to ω_{i,j_i^*} are independent of each other as well in $\widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha},\boldsymbol{\beta}^*)$. Therefore, we can directly draw an important conclusion that achieving the optimality of $\widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha},\boldsymbol{\beta}^*)$ is equivalent to achieving the optimality of each ω_{i,j_i^*}, where the optimality can be reached when $\boldsymbol{\alpha} = \boldsymbol{\alpha}^*$.

In view of the above, we know that $\boldsymbol{\alpha}_i^*$ must be the optimal solution of $\omega_{i,j_i^*}, \forall i \in \mathcal{V}$. Further combined with another fact that $\omega_{i,j}$ is the objective of $\mathbf{P1}_{i,j}, \forall (i,j) \in \mathcal{V}\times\mathcal{V}_i, j > i$ where $\boldsymbol{\alpha}_{i_{(j)}}^*$ is the corresponding optimal solution, we ensure that the equality $\boldsymbol{\alpha}_{i_{(j)}}^* = \boldsymbol{\alpha}_i^*$ holds when $j = j_i^*$.

Appendix B
Proof of Proposition 2

Suppose that $(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*)$ is not optimal for the problem in (12), which means there must exist another solution, denoted as $\bar{\boldsymbol{\alpha}} = \left[\bar{\boldsymbol{\alpha}}_1, \bar{\boldsymbol{\alpha}}_2, \cdots, \bar{\boldsymbol{\alpha}}_V\right]^T$ and $\bar{\boldsymbol{\beta}} = \left[\beta_{1\looparrowright\bar{j}_1}, \beta_{2\looparrowright\bar{j}_2}, \cdots, \beta_{V\looparrowright\bar{j}_V}\right]^T$, such that

$$\widetilde{L}_{\boldsymbol{\tau}}(\bar{\boldsymbol{\alpha}},\bar{\boldsymbol{\beta}}) < \widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*). \tag{30}$$

On the one hand, since $\boldsymbol{\beta}^*$ is the optimal solution to $\mathbf{P2}$, for $\bar{\boldsymbol{\beta}} \neq \boldsymbol{\beta}^*$, we have $\widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*) < \widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha}^*,\bar{\boldsymbol{\beta}})$. On the other hand, directly applying the conclusions in Proposition 1, it is seen that $\forall i \in \mathcal{V}$, we have $\boldsymbol{\alpha}_{i_{(j)}}^* = \bar{\boldsymbol{\alpha}}_i$ when $j = \bar{j}_i$. Combined with the previous assumption that $\bar{\boldsymbol{\beta}}$ is the optimal VSP solution to problem (12), $\omega_{i,\bar{j}_i}^* = \bar{\omega}_{i,\bar{j}_i}$ holds such that

$$\widetilde{L}_{\boldsymbol{\tau}}(\bar{\boldsymbol{\alpha}},\bar{\boldsymbol{\beta}}) = \widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha}^*,\bar{\boldsymbol{\beta}}) > \widetilde{L}_{\boldsymbol{\tau}}(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*). \tag{31}$$

However, there is a contradiction between (30) and (31). Consequently, the assumption cannot hold, which means that $(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*)$ is exactly the optimal solution to problem (12).

Acknowledgment

The authors would like to thank Dr. Rafiq Swash and Dr. Wallizada Mohibullah from AIDrivers Ltd. for their comments, which have substantially improved the quality of this paper.

References

[1] L. Xia, Y. Sun, D. Niyato, K. Ma, J. Kang, and M. A. Imran, "Knowledge base aware semantic communication in vehicular networks," in *IEEE ICC 2023 - IEEE International Conference on Communications (ICC)*, 2023.

[2] J. Wang, J. Liu, and N. Kato, "Networking and communications in autonomous driving: A survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1243–1274, 2018.

[3] C. She, C. Sun, Z. Gu, Y. Li, C. Yang, H. V. Poor, and B. Vucetic, "A tutorial on ultrareliable and low-latency communications in 6G: Integrating domain knowledge into deep learning," *Proceedings of the IEEE*, vol. 109, no. 3, pp. 204–246, 2021.

[4] W. Weaver, "Recent contributions to the mathematical theory of communication," *ETC: A Review of General Semantics*, pp. 261–281, 1953.

[5] J. Bao, P. Basu, M. Dean, C. Partridge, A. Swami, W. Leland, and J. A. Hendler, "Towards a theory of semantic communication," in *2011 IEEE Network Science Workshop*. IEEE, 2011, pp. 110–117.

[6] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, "Deep learning enabled semantic communication systems," *IEEE Transactions on Signal Processing*, vol. 69, pp. 2663–2675, 2021.

[7] E. C. Strinati and S. Barbarossa, "6G networks: Beyond Shannon towards semantic and goal-oriented communications," *Computer Networks*, vol. 190, p. 107930, 2021.

[8] G. Shi, Y. Xiao, Y. Li, and X. Xie, "From semantic communication to semantic-aware networking: Model, architecture, and open problems," *IEEE Communications Magazine*, vol. 59, no. 8, pp. 44–50, 2021.

[9] L. Xia, Y. Sun, X. Li, G. Feng, and M. A. Imran, "Wireless resource management in intelligent semantic communication networks," in *IEEE INFOCOM 2022 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2022, pp. 1–6.

[10] X. Luo, H.-H. Chen, and Q. Guo, "Semantic communications: Overview, open issues, and future research directions," *IEEE Wireless Communications*, pp. 1–10, 2022.

[11] L. Xia, Y. Sun, C. Liang, D. Feng, R. Cheng, Y. Yang, and M. A. Imran, "WiserVR: Semantic communication enabled wireless virtual reality delivery," *IEEE Wireless Communications*, vol. 30, no. 2, pp. 32–39, 2023.

[12] S. Gyawali, S. Xu, Y. Qian, and R. Q. Hu, "Challenges and solutions for cellular based V2X communications," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 222–255, 2020.

[13] M. Chein and M.-L. Mugnier, *Graph-based Knowledge Representation: Computational Foundations of Conceptual Graphs*. Springer Science & Business Media, 2008.

[14] H. Xie and Z. Qin, "A lite distributed semantic communication system for Internet of Things," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 1, pp. 142–153, 2020.

[15] L. Yan, Z. Qin, R. Zhang, Y. Li, and G. Y. Li, "Resource allocation for text semantic communications," *IEEE Wireless Communications Letters*, 2022.

[16] C. Guo, L. Liang, and G. Y. Li, "Resource allocation for vehicular communications with low latency and high reliability," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 3887–3902, 2019.

[17] S. T. Piantadosi, "Zipf's word frequency law in natural language: A critical review and future directions," *Psychonomic Bulletin & Review*, vol. 21, no. 5, pp. 1112–1130, 2014.

[18] G. Lavee, E. Rivlin, and M. Rudzsky, "Understanding video events: A survey of methods for automatic interpretation of semantic occurrences in video," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 39, no. 5, pp. 489–504, 2009.

[19] S. M. Ross, *Introduction to Probability Models*. Academic Press, 2014.

[20] F. Pollaczek, "Über eine Aufgabe der Wahrscheinlichkeitstheorie. I," *Mathematische Zeitschrift*, vol. 32, no. 1, pp. 64–100, 1930.

[21] H. Kellerer, U. Pferschy, and D. Pisinger, "Multidimensional knapsack problems," in *Knapsack Problems*. Springer, 2004, pp. 235–283.

[22] H. Tuy, T. Hoang, T. Hoang, V.-n. Mathématicien, T. Hoang, and V. Mathematician, *Convex Analysis and Global Optimization*. Springer, 1998.

[23] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. Courier Corporation, 1998.

[24] S. Burer and A. N. Letchford, "Non-convex mixed-integer nonlinear programming: A survey," *Surveys in Operations Research and Management Science*, vol. 17, no. 2, pp. 97–106, 2012.

[25] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[26] S. Boyd, L. Xiao, and A. Mutapcic, "Subgradient methods," *Lecture Notes of EE392o, Stanford University, Autumn Quarter*, vol. 2004, pp. 2004–2005, 2003.

[27] S. Salhi, "Defining tabu list size and aspiration criterion within tabu search methods," *Computers & Operations Research*, vol. 29, no. 1, pp. 67–86, 2002.

[28] F. Glover, "Tabu search-part I," *ORSA Journal on Computing*, vol. 1, no. 3, pp. 190–206, 1989.

[29] S. Diamond and S. Boyd, "CVXPY: A Python-embedded modeling language for convex optimization," *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 2909–2913, 2016.

[30] Y. T. Lee, Z. Song, and Q. Zhang, "Solving empirical risk minimization in the current matrix multiplication time," in *Conference on Learning Theory*. PMLR, 2019, pp. 2140–2157.

[31] Y. Song, K. W. Sung, and Y. Han, "Impact of packet arrivals on Wi-Fi and cellular system sharing unlicensed spectrum," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 10 204–10 208, 2016.

[32] *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Study on LTE-Based V2X Services; (Release 14)*, document TR 36.885, V2.0.0, 3GPP, Jun. 2016.

[33] L. Liang, G. Y. Li, and W. Xu, "Resource allocation for D2D-enabled vehicular communications," *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 3186–3197, 2017.

[34] R. Zhang, R. Lu, X. Cheng, N. Wang, and L. Yang, "A UAV-enabled data dissemination protocol with proactive caching and file sharing in V2X networks," *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 3930–3942, 2021.

[35] G. Ding, J. Yuan, G. Yu, and Y. Jiang, "Two-timescale resource management for ultrareliable and low-latency vehicular communications," *IEEE Transactions on Communications*, vol. 70, no. 5, pp. 3282–3294, 2022.

Le Xia (Graduate Student Member, IEEE) received the B.Eng. degree in communication engineering and the M.Eng. degree in electronics and communication engineering from the University of Electronic Science and Technology of China (UESTC) in 2017 and 2020, respectively. He is currently pursuing his Ph.D. degree with James Watt School of Engineering, the University of Glasgow, United Kingdom. His research interests include semantic communications, intelligent vehicular networks, and resource management in next-generation wireless networks.

Yao Sun (Senior Member, IEEE) is currently a Lecturer with James Watt School of Engineering, the University of Glasgow, Glasgow, UK. Dr Sun has won the IEEE Communication Society of TAOS Best Paper Award in 2019 ICC, IEEE IoT Journal Best Paper Award 2022 and Best Paper Award in 22nd ICCT. His research interests include intelligent wireless networking, semantic communications, blockchain system, and resource management in next generation mobile networks. Dr. Sun is a senior member of IEEE.

Dusit Niyato (M'09-SM'15-F'17, IEEE) is a professor in the School of Computer Science and Engineering, at Nanyang Technological University, Singapore. He received B.Eng. from King Mongkut's Institute of Technology Ladkrabang (KMITL), Thailand in 1999 and Ph.D. in Electrical and Computer Engineering from the University of Manitoba, Canada in 2008. His research interests are in the areas of sustainability, edge intelligence, decentralized machine learning, and incentive mechanism design.

Daquan Feng (Member, IEEE) is now an associate professor with the Shenzhen Key Laboratory of Digital Creative Technology, the Guangdong Province Engineering Laboratory for Digital Creative Technology, the Guangdong-Hong Kong Joint Laboratory for Big Data Imaging and Communication, College of Electronics and Information Engineering, Shenzhen University, Shenzhen, China. His research interests include URLLC communications, MEC, and massive IoT networks. Dr. Feng is an Associate Editor of IEEE Communications Letters, ICT Express, and Digital Communications and Networks.

Lei Feng (Member, IEEE) received his B.Eng. and Ph.D. degrees in Communication and Information Systems from the Beijing University of Posts and Telecommunications (BUPT) in 2009 and 2015, respectively. He is currently an Associate Professor at present in State Key Laboratory of Networking and Switching Technology, BUPT. His research interests include knowledge-driven next generation networks, 5G time-sensitive network technologies for Industrial IoT, and resource management and coordination of mobile networks.

Muhammad Ali Imran (Fellow, IEEE) received his M.Sc. (Distinction) and Ph.D. degrees from Imperial College London, UK, in 2002 and 2007, respectively. He is a Professor in Communication Systems in the University of Glasgow. He has a global collaborative research network spanning both academia and key industrial players in the field of wireless communications. He has supervised 50+ successful PhD graduates and published over 600 peer-reviewed research papers including more than 100 IEEE Transaction papers. Prof. Imran is a Fellow of IEEE.