Love of Variety based Latency Analysis for High Definition Map Updating: Age of Information and Distributional Robust Perspectives

Dawei Chen, *Member, IEEE*, Yifei Zhu, *Member, IEEE*, Dan Wang, *Senior Member, IEEE*, Haoxin Wang, *Member, IEEE*, Jiang Xie, *Fellow, IEEE*, Xiao-Ping Zhang, *Fellow, IEEE*, and Zhu Han *Fellow, IEEE*

***Abstract*—High definition (HD) map is a key technology that enables autonomous driving, which has the characteristic of frequent updates and low delay requirements. In order to minimize the HD map updating latency, this paper jointly considers the generation and transmission processes of HD map. The generation process of HD map relies heavily on the data captured by different types of sensors, which is modeled by the federated analytics process. Also, in order to quantify the diversity of utilized sensors, the concept of love of variety is introduced in this paper. Then, considering HD map is composed of dynamic and static layers that have different latency requirements during the transmission process, this paper proposes a method to allocate edge server capacity to each HD map layer such that the overall information staleness can be minimized. Firstly, the deterministic edge capacity case is discussed and the solution is derived by obtaining the Karush–Kuhn–Tucker conditions. Then, considering the practice that an edge server provides services to multiple attached devices simultaneously, contentions among these devices make available capacity for the autonomous vehicle variational. Therefore, an uncertain edge capacity case is discussed as well, where the uncertainty is described by Wasserstein metrics and the problem is reformulated into distributional robust chance constrained optimization problem. And for the low latency purpose, we utilize an inner approximation to reduce the complexity of the original problem and find a suboptimal solution. Finally, simulation results demonstrate the effectiveness of our proposed methods that achieves the lowest latency for HD map updating.**

***Index Terms*—high definition map, autonomous driving, federated analytics, love of variety, age of information, distributional robust chance constrained optimization.**

Dan Wang's work is supported by GRF 15210119, 15209220, 15200321, 15201322, CRF C5018-20G. Jiang Xie's work is supported in part by the US National Science Foundation (NSF) under Grant No. 1910667, 1910891, and 2025284. Xiao-Ping Zhang's work is supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) under Grant RGPIN-2020-04661. Zhu Han's work is partially supported by NSF CNS-2107216, CNS-2128368, CMMI-2222810, Toyota and Amazon.

Dawei Chen is with Toyota Motor North America R&D, InfoTech Labs, Mountain View, CA 94043 USA (email: dawei.chen1@toyota.com).

Yifei Zhu is with the University of Michigan–Shanghai Jiao Tong University Joint Institute, Shanghai Jiao Tong University, Shanghai 200240, China (email: yifei.zhu@sjtu.edu.cn).

Dan Wang is with the Department of Computing, Hong Kong Polytechnic University, Kowloon, Hong Kong (email: csdwang@comp.polyu.edu.hk).

Haoxin Wang is with the Department of Computer Science, Georgia State University, Atlanta, GA 30303 USA (e-mail: haoxinwang@gsu.edu).

Jiang Xie is with the Department of Electrical and Computer Engineering, University of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, NC 28223 (e-mail: linda.xie@uncc.edu).

Xiao-Ping Zhang is with the Department of Electrical, Computer and Biomedical Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada (e-mail: xzhang@ryerson.ca).

Zhu Han is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004 USA, and also with the Department of Computer Science and Engineering, Kyung Hee University, Seoul, 446-701, South Korea (email: zhan2@uh.edu).

I. Introduction

Having stepped into the era of information technology, there are enormous artificial intelligence based autonomous devices, technologies, and services coming into being, and one important branch of which is autonomous vehicles or intelligent vehicles. According to the definition of the National Highway Traffic Safety Administration (NHTSA), the levels of vehicle automation can be categorized into six classes, which are distinguished by the extent of autonomy [1]. Currently, the performance of autonomous vehicles can just meet the requirements between levels 2 and 3, and both of which require the driver must be ready to take back control at any time. Aimed at achieving a higher automation level, one effective way is to utilize the high-definition (HD) map. Unlike the traditional map, HD map is represented with a high degree of precision and resolution, which is as fine as 10-20 centimeters or better.

To generate and maintain HD map, one key is to extract useful information from data captured by embedded sensor systems, such as object detection, lane marking detection, ranging, etc. However, since different sensor systems have essential pros and cons regarding range, resolution, sensitivity to visibility, etc. [2]. In order to extract effective information that contributes to HD map generation and maintenance, collaborative making use of diverse sensors is inevitable, which exactly can be done by *federated analytics*. The aim of federated analytics is to obtain data insights among distributed devices by applying data science methods to the analysis of raw data generated on different clients [3]. Unlike federated learning, what transmitted between clients and aggregator are data insights rather than machine learning model gradients, such as distribution, positions of detected objects, etc. In other words, to support basic data science needs is the purpose of federated analytics.

In federated mechanisms, if the number of participants (i.e., sensors here) is larger, the desirable accuracy can be achieved faster and the optimization process converges more quickly. From the perspective of practice, this can be interpreted as well. When extracting information that contributes to HD map, if more types of sensors are utilized, the obtained information, (e.g. positions of detected objects), will be more accurate,

This article has been accepted for publication in IEEE Transactions on Intelligent Vehicles. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIV.2022.3224655

leading to a better accuracy level of HD map. For example, in a bad weather condition (e.g., fog, heavy rains, or storm), poor visibility will degrade the performance of visual-based sensors like LiDar and camera but has no influence on Radar and ultrasonic sensors [4]. In other words, the higher the diversity of utilized sensors is, the more accurate HD map can be generated. However, the challenge is how to quantify the impact of sensor diversity on the obtained HD map accuracy level.

To tackle this challenge, we model the diversity of sensors in the federated analytics problem into the utility of *love of variety*. Basically, love of variety is a concept from economics, which assumes that each consumer has a demand for multiple varieties of a product over a given time period [5]. Within the time period, the utility will increase if consumers use more different products. For HD map generation, the interpretation is that, during the federated analytics, the more types of sensor data are utilized, the better accuracy of HD map can be achieved. Since HD map is typically used for autonomous driving, the accuracy level has a required threshold, such that the probability of accidents can be reduced. Therefore, in order to achieve the desired accuracy level of HD map, if more types of sensors are included in federated analytics, the accuracy of each aggregation round will be higher. Thus, the total needed number of global aggregation rounds can be decreased.

Previously, the problem of HD map generation is discussed from one single vehicle's view. But from the practical use of HD map, uploading individual HD map to the server needs to be considered to form an intact city-wised HD map. Roughly, the information in the HD map can be categorized into the dynamic layer and the static layer. Dynamic layer information includes pedestrian, adjacent vehicles, and etc. While static layer information contains lanes, traffic signs, road markings, etc. Obviously, dynamic information updating is more latency-sensitive since it may vary within seconds [6], [7]. Fortunately, emerging multi-access edge computing (MEC) provides a low-latency paradigm for data transmission, which is realized by edge servers deployed at the edge of networks and the transmission distance can be reduced compared with traditional cloud computing [8], [9]. Because the different layer of HD map has a different requirement for updating latency. In order to quantify the latency and measure the freshness of HD map, the *age of information* (AoI) is employed here. AoI is an end-to-end metric that can be used to characterize latency in status updating systems and applications. In AoI, information freshness and staleness can be quantified through the penalty function. Different characteristics of penalty functions, like tendency, can describe the latency features of the specific application. For example, for HD map updating, the penalty will increase exponentially when the information staleness becomes larger as it is a latency-sensitive application and a lager latency may result in a critical result.

With different sensor diversity and desired accuracy level, the data size of the final obtained HD map based on federated analytics will be different. Therefore, consider the capacity of the edge server is limited, we propose a method to allocate the edge server capacity into different HD map layers so that the overall penalty can be minimized. Firstly, we discuss the case where the edge server capacity is deterministic. The optimal allocation scheme can be derived through Karush–Kuhn–Tucker (KKT) conditions. Then, considering the practice that an edge server provides services to multiple attached devices simultaneously, such that the available capacity for the autonomous vehicle is variational. We describe this ambiguity using Wasserstein metrics and reformulate the problem into a *distributional robust chance constrained optimization problem*, which can effectively and efficiently lead to a near-optimal solution.

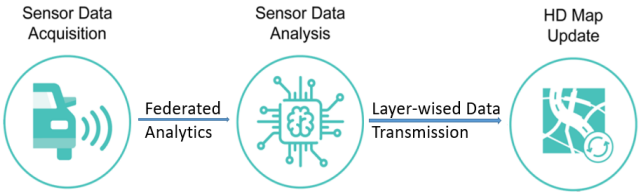


Fig. 1. System Overview. Firstly, the different kinds of on-vehicle sensors will capture the data. Then, these data will be analyzed in a federated analytics way to obtain insights. Finally, since latency demands for each layer in HD map are different, layer-wise data transmission will be performed to meet these distinguished requirements.

The overview of the system is illustrated in Fig. 1. The contribution of this work is summarized as follows:

- Because different sensors have their own pros and cons, it is inevitable to collaboratively use them to generate HD map. Therefore, we model the data analysis on various types of sensors into a federated analytics process. In order to quantify the influence of diversity of utilized sensors among HD map accuracy, we use the love of variety approach to evaluate the diversity.
- Practically, HD map is made up of dynamic and static information, which has different requirements for latency. To model this, we introduce the penalty function of AoI to quantify this disparity.
- Considering different latency demands for each layer in HD map, we propose a method to allocate transmission bandwidth to each HD map layer such that the overall information staleness can be minimized. Moreover, deterministic and uncertain edge capacity cases are both discussed in this work.
- We conducted the simulations for both deterministic and uncertain capacity cases. The results show that our proposed methods can effectively reduce the latency for HD map updating.

The remainder of this article is organized as follows. Section II discusses some related existing literature. Section III introduces the details of HD map composition and sensor systems in autonomous vehicles. Section IV gives the system model for federated analytics and HD map layer-wised transmission and formulates the corresponding information staleness minimization problem for HD map updating. In Section V, the proposed allocation algorithm for the deterministic capacity case is introduced. Section VI first introduces the mathematical interpretation for uncertainty and then gives a solution to the uncertain capacity case. Section VII conducts the simulation

experiments and interprets the results accordingly. Finally, the conclusion is drawn in Section VIII.

II. Related Work

As a key technique to realize autonomous driving, HD map has attracted interest from both industry and academia. Some existing literature focuses on how to make use of HD map. [10] uses near-term future information of HD map to propose a control scheme enabling predictive cruise control such that the overall fuel consumption can be minimized. [11] proposes an image region of interest extraction method to improve the accuracy of traffic light recognition with the help of HD map and self-localization techniques. [12] utilizes HD map as prior information, and proposes a local motion planning method to realize path planning and obstacles avoidance in autonomous driving scenarios. [13] proposes a sequential algorithm that can perform accurate lane-keeping or changing decisions while keeping a safe distance from the adjacent vehicle by using the position of the host vehicle and HD map. However, these works concentrate on HD map enabled applications instead of the generation and updating strategies.

While some of the existing literature focuses on the HD map transmission. [14] proposes a collaborative vehicle to everything (V2X) transmission scheme to meet the transmission rate requirement for HD map while achieving low power consumption. [15] proposes a distributed multi-agent multi-armed bandit algorithm to maximize the accumulated cache utility for each roadside unit (RSU) by caching appropriate HD map contents in storage. [16] proposes a cluster based strategy for HD map offloading by using characteristics of HD map data and mobility of vehicles so that the energy consumption and offloading delay can be minimized. [17] proposes a fusing algorithm based on the Kalman filter to increase the position and semantic confidence of HD map such that the efficiency of HD map updating can be improved. [18] proposes a HD map data distribution mechanism such that the HD map provision task can be allocated to the selected RSU and transmit proportionate HD map data for energy efficiency purposes.

Also, there is some work focusing on the creation of HD map. [19] introduces the workflow of HD map creation and the machine learning based techniques used by industry that can minimize the amount of manual work for HD map generation. [20] proposes a crowd-sourcing framework to update the point cloud map layer in HD map from environment changes by jointly utilizing LiDar and vehicle communication. [21] proposes a semantic-based road segmentation method to address the problems of dynamic obstacles and shadows as well as the GNSS signal errors for HD map construction. [22] proposes a Dislocation-Projection approach to create a color-pointed layer of HD map based on effective processing of LiDar, camera, and global navigation satellite system (GNSS) data. However, all of the above literature only focuses on one perspective of HD map, neither creation, transmission, nor application. While our work jointly considers the generation and transmission processes of HD map in a layer-wised and latency-aware manner. Besides, the generation of HD map needs contributions from different types of sensors, which is regarded as a federated analytics problem and almost no existing literature does this way.

III. Preliminaries

In this section, the background information is introduced. The basics and layered HD map are introduced in Subsection III-A, the sensor systems are discussed in Subsection III-B.

A. High Definition Map

HD map is made up of various information and resources, such as drivable paths, lane marks, the priority of lanes, traffic light and crosswalk to lane association, adjacent objects, and street furniture, which is represented in a high degree of resolution and precision, generally in the centimeter level. For practical autonomous driving use cases, HD map is the indispensable key for the advanced driver assistance system (ADAS). Intuitively, the contents of HD map can be roughly categorized into two classes: dynamic objects (such as pedestrians and vehicles) and static objects (such as traffic signs and lights). According to the definition of automotive edge computing consortium (AECC), the composition of HD map is layered, which can be represented by the highly dynamic layer, transient dynamic layer, transient static layer, and permanent static layer, as is illustrated in upper part of Fig. 2 [23].

- In the highly dynamic layer, contents include the position, velocity, and accelerator of pedestrians, vehicles, and so on, which changes in several seconds.
- In the transient dynamic layer, contents include obstacles like fallen objects and trash, and local weather like unexpected sudden rain, which changes in several minutes.
- In the transient static layer, contents include road work, temporary road closure, accidents, and so on, which changes in several hours.
- In the highly static layer, contents include lanes, traffic lights, road signs, and so on, which change in several days or longer.

In the current map, it is common to find the gaps between actual circumstances and the map, which takes days to be corrected [23]. However, for autonomous driving, any delayed update for HD map can result in dangerous or even fatal accidents without human intervention. Therefore, HD map must be updated in a timely manner.

B. Sensor Systems in Autonomous Vehicles

HD map is inevitable for autonomous driving. How to generate and maintain this map with abundant dynamic and static information is challenging. Basically, HD map is mainly created by on-board sensor data. Typically, there are three main sensors in automotive vehicles in the present: cameras, Radar (Radio Detection And Ranging), and LiDar (Light Detection And Ranging). They allow the vehicle to see and sense everything on the road, as well as to collect the information needed in order to drive safely.

- *Camera*: Autonomous cars often have several smart cameras deployed in front and rear of the vehicle so that a

This article has been accepted for publication in IEEE Transactions on Intelligent Vehicles. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIV.2022.3224655

TABLE I
PROS AND CONS OF SENSOR SYSTEMS IN AUTONOMOUS VEHICLES

Sensor	Strength	Weakness
Radar	Good ranging accuracy and does not rely on visibility	Cannot detect road markings
LiDar	Highly accurate ranging	Higher cost and less effective in featureless areas (like country roads)
Camera	Good object detection Good in tunnels/ urban canyons	Less effective in featureless roads, low visibility scenario (snow, rain, darkness, fog), roads without lane markings, and construction areas
GNSS	Good accuracy, Common global reference between vehicles, Good in low visibility, featureless roads	Less reliable in high urban canyons
IMU	Good in low visibility, featureless roads	Not usable in long tunnels (a few km) due to high drift rate of IMU
Ultrasonic	Good in low visibility, featureless roads	Need close proximity and slow speeds

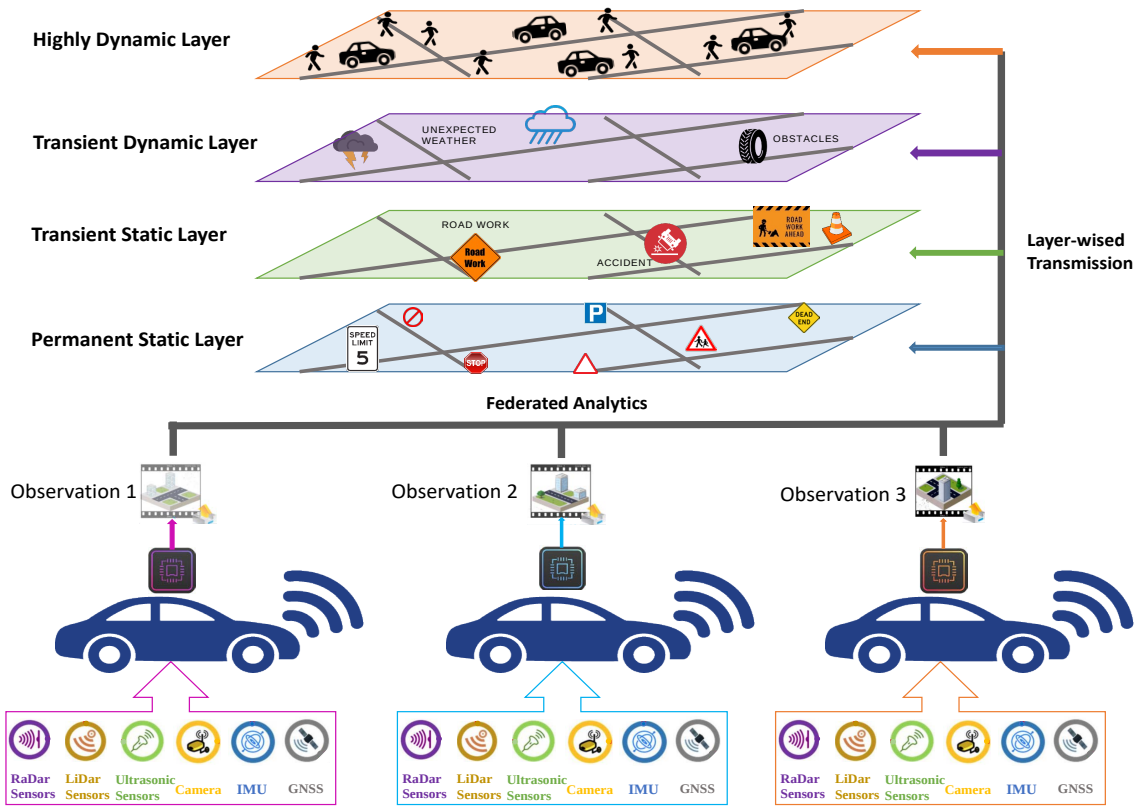


Fig. 2. Illustration of federated analytics for HD map generation. HD map is mainly created by on-board sensor data, the lower part of the figure shows the sensor data harvesting process. Since different sensor systems have essential pros and cons, and they are complementary to each other. Processing these different sensor data can be described by federated analytics. Besides, since different vehicles have different locations and position, the obtained data are actually captured within a local area due to the limitaion of sensing range. Therefore, to generate a complete HD map, crowdsourcing data from multiple vehicles are necessary, which is illustrated in the middle part of the figure. Then, since latency demands for each layer in HD map is different, a layer-wise data transmission will be performed to meet these distinguished requirements, which is illustrated in the upper part of the figure.

360° view of the external environment can be generated [24]. Unfortunately, these camera sensors are still far from perfect. Poor weather conditions such as rain, fog, or snow can prevent cameras from clearly seeing the obstacles in the roadway.

- *Radar*: Radar sensors send out radio waves that detect objects and gauge their distance and speed in relation to the vehicle in real time. Unlike camera sensors, radar systems typically have no trouble at all when identifying objects during fog or rain. But since the sensors only scan horizontally, which can cause a variety of problems when driving under bridges or canyons [25].
- *LiDar*: LiDar sensors work similar to radar systems, with the only difference being that they use lasers instead of radio waves. Apart from measuring the distances to various objects on the road, LiDar allows creating 3D images of the detected objects and mapping the surroundings [26]. The main problem for performance is the same as camera, i.e., poor weather and visibility can sometimes block LiDar sensors.

Except for the three kinds of sensors, there are also some traditional sensor systems, such as

- *GNSS*: GNSS is the technology that uses satellites constellation to provide autonomous geo-spatial positioning, navigation, and timing services on a global or regional

TABLE II
PARAMETER AND SYMBOL DESCRIPTION

Notation	Definition
$\mathcal{N}$	Total time slots index set.
α	Utility function parameter.
β	Utility function parameter.
ρ	Utility function parameter.
ϵ	Elasticity.
r_u	Relative love of variety for utility function u.
ψ	Relative accuracy level.
G	The number of global iterations.
ζ	Parameter for G.
k	Constant parameter.
L_i	The layer i in HD map.
r_i	Sensing range of layer i.
D_i	Data size of layer i.
B	Communication bandwidth.
w	Transmission power.
h	Channel gain.
N_0	Gaussian noise.
ϕ_i	Offloading part percentage for layer i.
Δ_i	Transmission latency for layer i.
p_i	Penalty of layer i.
γ_i	Penalty function parameter of layer i.

basis, which includes the GPS, GLONASS, Galileo, Beidou and other regional systems.

- *Inertial motion units (IMU)*: IMU is a device including multi-axea, accelerometers, and gyroscopes that can provide an estimation of an object movement in space, such as measures force, angular rate, attitude, and orientation.
- *Ultrasonic sensors*: Ultrasonic sensors mimic echolocation used by bats and are able to calculate distance between objects within short range through transmitting high-frequency sound waves.

The pros and cons of different sensor systems are summarized in Table. I.

IV. System Model and Problem Formulation

In this section, the concept of love of variety is introduced in Subsection IV-A. We model the utilization of multiple types of sensors for federated analytics into a love of variety based utility in Subsection IV-B. Then, the models of HD map layers and layer-wised transmission are introduced in Subsections IV-C and IV-D, respectively. The transmission time based AoI is introduced in Subsection IV-E. Finally, since different layers of HD map has different delay requirement, we formulate the HD map transmission penalty minimization problem from the perspective of AoI in Subsection IV-F. For a clear understanding of parameters and symbols in this paper, the main definitions and descriptions are provided in Table II in details.

A. Love of Variety

Obviously, when performing federated analytics, as illustrated in Fig. 2, the results will be better if more types of sensors are included, e.g., the object detection and ranging can be accurate in both good visibility and poor weather scenarios. We assume the electronic control unit (ECU) can perform one type of data in a specific time slot, and a vehicle has a demand for multiple varieties of sensor data over time. Therefore, a

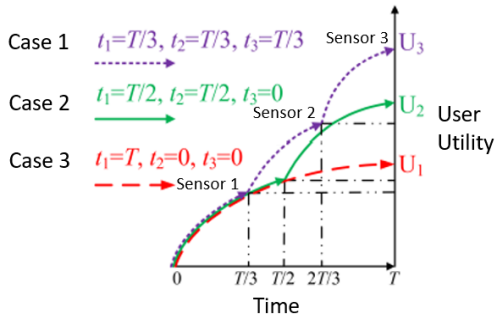


Fig. 3. Illustration of RLV.

time vector $\boldsymbol{t} = t_{i\in\mathcal{N}}$ can be used to describe the computation over different sensor systems, where t_i is denoted as ECU is executing a specific kind of sensor data during a fixed time slot, i is the time slot index, and $\mathcal{N}$ is the total time slots index set. Apart from the utilization of diverse sensors, the utility of a vehicle is also related to computation time slot t_i. Therefore, the utility function of ECU can be defined as $u(t_i)$, which is a strictly increasing and concave function and meets $u(0) = 0$, as is suggested in [27]. The general utility function can be defined as

$$u(t_i) = \frac{1}{1-\rho}\left[(\alpha + t_i)^{1-\rho} - \alpha^{1-\rho}\right] + \beta t_i, \tag{1}$$

where $\alpha \geq 0$, $\beta \geq 0$, and $0 < \rho < 1$ are constant coefficients. Intuitively, the utilities from computation over different types of sensors are additive. Overall, the aggregated utility of an ECU or vehicle is $\sum_{i\in\mathcal{N}} u(t_i)$.

Mathematically, to involve variety, one key challenge is how to evaluate or quantify the computation ability of an ECU to hand over computation from sensor data x to another sensor data y, or how to quantify the diversity of sensor data computed within a certain time period. Besides, since each vehicle needs diverse sensor data to obtain more accurate and effective information for HD map, how to quantify willingness of exchanging among multiple sensors is also challenging. In order to solve this problem, we introduce elasticity, whose definition is shown below.

***Definition** 1:* For two variables x and y, the x-elasticity of y is defined as

$$\epsilon_x^y = -\frac{\partial y}{\partial x}\frac{x}{y}. \tag{2}$$

The interpretation of elasticity is that the percentage change in y is in response to the percentage change in x. If the value of elasticity is larger, it means y is more sensitive to the change of x. To quantify the willingness of exchanging among multiple sensors, the definition of relative love of variety (RLV) is given as the following.

***Definition** 2:* The vehicle's relative love of variety is the elasticity of the marginal utility with respect to the computation time slot t_i, which is described by

$$r_u(t_i) = \epsilon_{t_i}^{u'} = -\frac{u'' t_i}{u'} > 0. \tag{3}$$

Obviously, from Definition 2, the value of RLV reflects whether the vehicle is willing to exchange different sensor

data in consecutive time slots for achieving a higher marginal utility, as is described in Fig. 3. For case 3, the user changes consumption among sensor data at the end of time slot $T/3$ and achieves the highest utility U_3. For case 2, the user consumes two types of sensor data within time T and achieves utility U_2. While the user in case 1 keeps the same type of sensor data throughout time interval and achieves the lowest utility U_1. Besides, it should be noted that the definition of RLV is with respect to a particular sensor generated data. When we need to analyze the overall RLV difference among multiple vehicles, we can utilize the average or summation of RLVs from all the types of sensor data to describe the overall RLV level, i.e., $\bar{U} = \frac{1}{N}\sum_{i=1}^{N} u(t_i)$.

B. Data Quality for Federated Analytics

Obviously, when performing federated analytics, the results will be better if more vehicles with high love of variety unity are included [28]. Since a higher love of variety utility means the utilization of number of different sensors is larger. As is discussed in Subsection III-B, the functionalities of sensors are complementary to each other. Therefore, higher love of variety utility means a more accurate HD map can be obtained or the data quality is better. So we can utilize the utility value or average RLV to describe the precision or accuracy. According to [29], for distributed optimization, if the global optimization problem is strongly convex, the general upper bound on number of global iteration is $\frac{O(log(\frac{1}{\psi}))}{1-\psi}$, where ψ is relative accuracy level of the local subproblem. Then, the needed number of global iterations can be given as

$$G = \frac{\zeta \cdot \log(\frac{1}{\psi})}{1-\psi}, \tag{4}$$

where $\zeta > 0$ is a constant. It can be seen that for a certain number of global iterations, the global accuracy level can be improved (i.e., ψ is close to 0) solving local subproblems towards high accuracy. However, the inverse dependence on $1-\psi$ means that there is a limit to how much the global accuracy can gain from the high-accurate solutions of the local subproblems [30]. To obtain ψ-accuracy, the global iteration will always require $\zeta \cdot \log(\frac{1}{\psi})$ [31].

For federated mechanisms, if the number of participants is larger, higher accuracy can be achieved when the number of global aggregation is fixed. Likewise, here, when the number of types of sensor is larger, the value of RLV or the utility U will be larger accordingly. When a desirable accuracy level ψ is given, the needed number of global aggregation will be less compared with the low average variety vehicle group. Therefore, for the vehicles with average utility level $\bar{U}$, the needed number of global federated analytics aggregation will be

$$S = \kappa G \left\lceil \frac{\log(\frac{1}{\psi})}{\bar{U}} \right\rceil, \tag{5}$$

where κ is a constant such that the given accuracy level ψ can be achieved. Obviously, if the average utility is larger, the total number of iteration will be less when the global accuracy level is fixed, as is proved in [28].

C. HD Map Components

Let $L_1(r_1)$, $L_2(r_2)$, $L_3(r_3)$, and $L_4(r_4)$ denote highly dynamic layer, transient dynamic layer, transient layer, and permanent static layer, respectively. r_i is the range that needs to be updated. Here, r_i for each layer is different, which means each HD map layer has an essential requirement for updating range. For example, for the permanent static layer, the required info may be "on street A (like several miles long or more), the speed limit is 50mph". However, for the highly dynamic layer, the vehicle needs to know the nearest pedestrian within several meters or less, which is much more precise than the permanent static layer requirement. Therefore, the precision scale from the permanent static layer, transient static layer, transient dynamic layer, to the highly dynamic layer can be county wide, region wide, a few blocks wide, to a crossing wide. Therefore, for each time updating, i.e., performing a federated analytics global aggregation, the size of each layer can be represented by

$$D_i = k \times r_i, \tag{6}$$

where k is a constant.

Moreover, in order to achieve ψ-accuracy, the total number of the HD map size will be

$$D = \sum_{i=1}^{4} D_i = \sum_{i=1}^{4} Skr_i. \tag{7}$$

Here, the number of components is four, which is because the number of layers to compose HD map is four, as discussed in Subsection III-A.

D. Communication Model

Intuitively, since we have different requirements (scale and time) for different layers. For those with no reporting of the real-time object layers, e.g. the transient static layer and permanent static layer, usually with relatively larger required areas, the updating can be done via the cloud through the backhaul network. While for those reporting of real-time objects layers, e.g. the transient dynamic layer and highly dynamic layer, usually with relatively smaller required regions, the HD map updating can be done through V2V communication or edge servers. For the updating part that is transmitted through the edge server, the transmission rate can be described as

$$v = B\log_2\left(1 + \frac{wh}{N_0}\right), \tag{8}$$

where B is the bandwidth, w is transmission power, h is channel gain, and N_0 is the Gaussian noise. For the updating that is done by backhaul network, we assume the delay to transmit a unit size map is a constant t_c, which is typically larger than t_e.

To update a certain layer, we assume the part transmitted via edge server is denoted by $\phi_i \in [0, 1]$. So in a time period T, the total time consumption is

$$t(\phi_i) = \sum_{i=1}^{4}\left(\frac{D_i\phi_i}{v} + D_i(1-\phi_i)t_c\right). \tag{9}$$

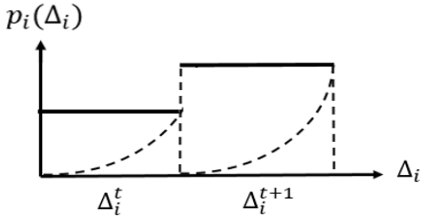


Fig. 4. Penalty function.

E. Staleness of Information Model

AoI is an end-to-end metric that can be used to characterize latency in status updating systems and applications [32]. An update packet with timestamp a is said to have age $b-a$ at a time $b \geqslant a$. Here, timestamp a denotes the generation time of HD map and timestamp b denotes the reception of updating [33]. Therefore, the difference $b-a$ describes the transmission time of HD map, i.e., $b-a=t(\phi_i)$. For each composition in HD map, the updating age Δ_i can be written as

$$\Delta_i = t(\phi_i). \tag{10}$$

Since the delay requirement for HD map updating is very strict, when the delay becomes larger in a certain time period, the penalty increases non-linearly at the same time. Therefore, the staleness of information updating can be defined as

$$p_i\left(t_i(\phi_i)\right) = t_i(\phi_i)^{\gamma_i}, \tag{11}$$

where γ_i is the constant and can describe sensitivity of different HD map layers regarding to delay. The typical cure of staleness is shown in Fig. 4.

F. Problem Formulation

For an autonomous vehicle, in order to minimize the time staleness for HD map updating, the problem can be formulated as

$$\begin{aligned} \min_{\phi_i} \quad & \sum_{i=1}^{4} p_i \\ s.t. \quad & \sum_{i=1}^{4} D_i\phi_i \leq q, \\ & \phi_i \geqslant 0, \end{aligned} \tag{12}$$

where q denotes the available computation or storage resources of edge server. Here, the constrain means the total edge transmission data of four layers cannot exceed the capacity of the edge server.

V. Deterministic Capacity Case

In problem (12), the staleness of each layer is p_i, which can be rewritten as

$$\begin{aligned} p_i &= \left[\frac{D_i\phi_i}{v} + D_i(1-\phi_i)t_c\right]^{\alpha_i} \\ &= \left[\left(\frac{1}{v} - t_c\right)D_i\phi_i + D_i t_c\right]^{\alpha_i} \\ &= \left(a_i x_i + b_i\right)^{\alpha_i}, \end{aligned} \tag{13}$$

where $x_i = D_i\phi_i$, $a_i = \left(\frac{D_i}{v} - D_i t_c\right)$ and $b_i = D_i t_c$ are layer characteristic constants. Therefore, problem (12) can be rewritten as

$$\begin{aligned} \min_{x_i} \quad & \sum_{i=1}^{4} \left(a_i x_i + b_i\right)^{\alpha_i} \\ s.t. \quad & \sum_{i=1}^{4} x_i \leq q, \\ & x_i \geqslant 0. \end{aligned} \tag{14}$$

To address problem (15), we have the following theorem.

***Theorem** 1:* Let x^* be a feasible point of

$$\begin{aligned} \min \quad & f(x) \\ s.t. \quad & g_i(x) \leq 0, \quad i = 1, \cdots, m, \\ & h_i(x) = 0, \quad i = 1, \cdots, n, \end{aligned} \tag{15}$$

where f and g_i are continuously differentiable convex functions over $\mathbb{R}$, and h_i are affine functions. Suppose that there exsist multipliers $\lambda_1, \cdots, \lambda_m$ and $\mu_1, \cdots, \mu_n \in \mathbb{R}$ such that

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{n} \mu_i \nabla h_i(x^*) = 0 \tag{16}$$

and

$$\lambda_i g_i(x^*) = 0, \quad i = 1, \cdots, m. \tag{17}$$

Then x^* is an optimal solution of (16).

proof 1: Let x be a feasible point. Define the convex function

$$L(x) = f(x) + \sum_{i}^{m} \lambda_i g_i(x) + \sum_{i=1}^{n} \mu_i h_i(x). \tag{18}$$

A feasible point x^* is a minimizer of L, since $\nabla L(x^*) = 0$ and in particular $L(x^*) < L(x)$, we have

$$\begin{aligned} f(x^*) &= f(x^*) + \sum_{i}^{m} \lambda_i g_i(x^*) + \sum_{i=1}^{n} \mu_i h_i(x^*) = L(x^*) \\ &\leq L(x) = f(x) + \sum_{i}^{m} \lambda_i g_i(x) + \sum_{i=1}^{n} \mu_i h_i(x) \\ &\leq f(x), \end{aligned} \tag{19}$$

showing that x^* is the optimal solution. ■

To find the solution to problem (14), we introduce Lagrange multiplier $\lambda^* \in \mathbb{R}$ for the inequality constraint $x_i \geqslant 0$, and a multiplier $\mu^* \in \mathbb{R}$ for the equality constraint $\sum x_i = q$, i.e.,

$$L = \sum \left(a_i x_i + b_i\right)^{\alpha_i} - \sum \lambda_i x_i + \mu\left(\sum x_i - q\right). \tag{20}$$

Then, we can obtain the KKT conditions. Firstly, we have the following lemma.

***Lemma** 1:* Let x^* and any (λ^*, μ^*) be any primal and dual optimal points with zero duality gap. Since x^* minimizes $L(x, \lambda^*, \mu^*)$ over x, it follows that its gradient must vanish at x^*.

Therefore, we have the stationarity condition, which is

$$a_i\alpha_i\left(a_i x_i^* + b_i\right)^{\alpha_i - 1} - \lambda_i^* + \mu^* = 0. \tag{21}$$

The complementary slackness is

$$\lambda_i^* x_i^* = 0. \tag{22}$$

Primal feasibility is

$$x_i^* \geqslant 0, \tag{23}$$

and

$$\sum x_i^* = q. \tag{24}$$

And dual feasibility is

$$\lambda^* \geqslant 0. \tag{25}$$

We can directly solve these equations to find x_i^*, λ^*, and μ^*. We start by noting that λ^* acts as a slack variable in (17), so it can be eliminated. So we have

$$\left(\mu^* + a_i\alpha_i(a_i x_i^* + b_i)^{\alpha_i - 1}\right) x_i^* = 0, \tag{26}$$

and

$$\mu^* \geqslant -a_i\alpha_i(a_i x_i^* + b_i)^{\alpha_i - 1}. \tag{27}$$

If $\mu^* < -a_i\alpha_i b_i^{\alpha_i - 1}$, condition (22) can only hold if $x_i^* > 0$. Solving for x_i^*, we conclude that $x_i^* = \frac{\sqrt[\alpha_i - 1]{\frac{\mu^*}{a_i\alpha_i} - b_i}}{a_i}$. If $\mu^* > -a_i\alpha_i b_i^{\alpha_i - 1}$, then $x_i^* > 0$ is impossible, because it would imply $\mu^* \geqslant -a_i\alpha_i b_i^{\alpha_i - 1} > -a_i\alpha_i(a_i x_i^* + b_i)^{\alpha_i - 1}$, which violates the complementary slackness condition. Therefore, $x_i^* = 0$ if $\mu^* \geqslant -a_i\alpha_i b_i^{\alpha_i - 1}$. Thus we have

$$x_i^* = \begin{cases} \frac{\sqrt[\alpha_i - 1]{\frac{\mu^*}{a_i\alpha_i} - b_i}}{a_i} & \mu^* < -a_i\alpha_i b_i^{\alpha_i - 1}, \\ 0 & \mu^* \geqslant -a_i\alpha_i b_i^{\alpha_i - 1}. \end{cases} \tag{28}$$

Alternatively, we can put in a more simple way,

$$x_i^* = \max\left\{0, \frac{\sqrt[\alpha_i - 1]{\frac{\mu^*}{a_i\alpha_i} - b_i}}{a_i}\right\}. \tag{29}$$

Substituting the expression for x_i^* into the equality constraint, we have

$$\sum_{i=1}^{4} \max\left\{0, \frac{\sqrt[\alpha_i - 1]{\frac{\mu^*}{a_i\alpha_i} - b_i}}{a_i}\right\} = q. \tag{30}$$

VI. Uncertain capacity case

The previous section discusses the solution to the deterministic edge capacity case. However, practically, an edge server provides services to multiple attached devices simultaneously, such that the available capacity q for the autonomous vehicle is variational. Therefore, problem (12) can be rewritten as

$$\begin{aligned} \min_{\phi_i} \quad & \sum_{j=1}^{4} p_i \\ s.t. \quad & \sum_{i=1}^{4} D_i\phi_i \leq \Delta q, \\ & \phi_i \geqslant 0, \end{aligned} \tag{31}$$

where Δq denotes the uncertainty of available capacity. In this section, the uncertain capacity allocation case is discussed. Subsection VI-A introduces the method using the Wasserstein distance to mathematically describe the ambiguity. Then, a distributionally robust chance constrained optimization based method is introduced to solve the uncertain capacity allocation problem in Subsection VI-B.

A. Wasserstein Distance Based Ambiguity Set

Assuming ambiguity set Δq follows a distribution, a natural way to hedge against the distributional ambiguity is to consider a neighborhood of the empirical probability distribution. Considering the discrepancy-based ambiguity sets: we introduce ambiguity sets based on probability distance:

$$\mathcal{P} = \{P : d(\widehat{P_N}, P) \leq \epsilon\}, \tag{32}$$

where $\hat{P_N}$ is empirical probability for capacity q distribution, ϵ is radius. $d(\hat{P_N}, P)$ is the metric to measure the similarity or distance of two distributions, and one of the most commonly used methods is the Wasserstein distance, which is defined as follows.

***Definition** 3:* The Wasserstein metric d_W: $\mathcal{M}(\Xi)\times\mathcal{M}(\Xi) \to \mathbb{R}$ is defined by

$$d_W(\mathbb{Q}_1, \mathbb{Q}_2) := \inf \int_{\Xi^2} \|\xi_1 - \xi_2\| \Pi(d\xi_1, d\xi_2), \tag{33}$$

for all distributions $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathcal{M}(\Xi)$, where $\|\cdot\|$ represents an arbitrary norm on $\mathbb{R}^m$, $\Pi(\mathbb{Q}_1, \mathbb{Q}_2)$ is the the set of all possible joint distributions of $\mathbb{Q}_1$ and $\mathbb{Q}_2$. ξ_1 and ξ_2 are the samples under joint distribution Π.

With the Wasserstein metric, the Wasserstein distance based ambiguity set can be defined as

$$\mathbb{B}_\epsilon(\widehat{P_N}) = \left\{\mathbb{Q} \in \mathcal{M}(\Xi) : d_W(\widehat{P_N}, \mathbb{Q}) < \epsilon\right\}. \tag{34}$$

The ambiguity set Q can be viewed as a Wasserstein ball which contains all probability distributions whose Wasserstein distance to the empirical distribution $\widehat{P_N}$ is less than ϵ [34]. The Wasserstein ball contains all possible probability distributions. $\mathbb{Q}$ will include the true distribution with a higher probability. Under a common light tail assumption on the unknown data-generating distribution, this ambiguity set offers attractive performance.

B. Distributional Robust Chance Constrained Optimization

With the Wasserstein metric, problem (13) can be reformulated as the following Distributional Robust Chance Constrained Optimization (DRCCO) form:

$$\begin{aligned} \min_{x} \quad & cx \\ s.t. \quad & x_i \in S, \\ & \inf_{P\in\mathcal{P}} P\left\{\tilde{\xi} : \tilde{\xi} \leq b(x)\right\} \geqslant 1 - \epsilon, \end{aligned} \tag{35}$$

where vector $x \in \mathbb{R}^n$ denotes the decision variables, vector $x \in \mathbb{R}^n$ is the objective function coefficients, set $S \subseteq \mathbb{S}^n$ is the deterministic constraints on x, and the last constraint is a chance constraint specified by the ambiguity set. Therefore, the second constraint is also named the distributionally robust chance constraint (DRCC). DRCC requires all of the uncertain constraints satisfied for all the probability distributions from ambiguity set P with a probability at least $(1 - \epsilon)$, where $\epsilon \in (0, 1)$ is the specified risk tolerance [35]. The feasible region induced by DRCC is defined as

$$Z := \left\{x \in \mathbb{R}^n : \inf_{P\in\mathcal{P}} P\{\tilde{\xi} : \tilde{\xi} \leq b(x)\} \geqslant 1 - \epsilon\right\}. \tag{36}$$

Firstly, by using the strong duality result from [36], (36) can be reformulated by the following theorem.

***Theorem** 2:* Set Z is equivalent to

$$Z = \left\{ x \in \mathbb{R}^n : \begin{array}{l} \delta - \epsilon\gamma \leq \dfrac{1}{N} \displaystyle\sum_{j\in[N]} z_j, \\ z_j + \gamma \leq \max\left\{b(x) - \zeta^j, \forall j \in [N], \right\}, \\ z_j \leq 0, \forall j \subseteq [N], \gamma \geqslant 0. \end{array} \right\}. \tag{37}$$

Further, we can reformulate set Z in (37) as a mixed integer program as below [37].

***Lemma** 2:* For DRCCO with right-hand uncertainty, suppose that there exists an $M \in \mathbb{R}_+^N$ such that

$$\max_{i\in[I]} \max_{x\in Z} \left\{|b(x) - \zeta^j|\right\} \leq M_j, \tag{38}$$

for $\forall j \in [M]$. Then set Z is mixed integer representable, i.e.,

$$Z = \left\{ x \in \mathbb{R}^n : \begin{array}{l} \delta - \epsilon\gamma \leq \dfrac{1}{N} \displaystyle\sum_{j\in[N]} z_j, \\ z_j + \gamma \leq s_j, \forall j \in [N] \\ s_j \leq b(x) - \zeta^j + M_j(1 - y_j), \forall j \in [N], \\ s_j \leq M_j y_j, \forall j \in [N], \\ \gamma \geqslant 0, z_j \leq 0, s_j \geqslant 0, y_i \in \{0,1\}, \forall j \in [N]. \end{array} \right\}. \tag{39}$$

However, the scale of programming (39) is large, which is difficult to meet the real-time requirements of HD map updating. Therefore, to quickly find a feasible solution needs to be considered. According to [38], the feasible set Z can be reformulated as descriptions by a conditional-value-at-risk (CVaR) constrained set. Regarding a random variable $\tilde{X}$, the $(1-\epsilon)$-value at risk (VaR) of $\tilde{X}$ is

$$\mathrm{VaR}_{1-\epsilon}(\tilde{X}) = \min\left\{z : F_{\tilde{X}} \geqslant 1 - \epsilon\right\}, \tag{40}$$

where F is the cumulative distribution function of $\tilde{X}$, defined by $F_{\tilde{X}} = P\{\tilde{X} \leq z\}$. And the $(1-\epsilon)$-CVaR is defined as

$$\mathrm{CVaR}_{1-\epsilon}(\tilde{X}) = \min_{\beta}\left\{\beta + \frac{1}{\epsilon}\mathbb{E}_P\left[\tilde{X} - \beta\right]_+\right\}. \tag{41}$$

It is observed that, for any random variable $\tilde{X}$, we have

$$\mathrm{CVaR}_{1-\epsilon}(\tilde{X}) \leq \mathrm{CVaR}_1(\tilde{X}) := \mathrm{ess.sup}(\tilde{X}). \tag{42}$$

Then the feasible solution can be approximated by the following theorem.

***Theorem** 3:* Set Z can be inner approximated by

$$Z_R = \left\{x \in \mathbb{R}^n : \frac{\delta}{\epsilon} + \zeta^j \leq b(x), \forall j \in [N]\right\}. \tag{43}$$

proof 2: Because $\mathrm{CVaR}_{1-\epsilon}\left[-f(x,\zeta)\right] \leq \mathrm{ess.sup}\left[-f(x,\zeta)\right]$, where ζ is a random vector, the feasible set Z can be inner approximated as

$$Z_R = \left\{x \in \mathbb{R}^n : P_\zeta\left\{f(x,\zeta) \geqslant \frac{\delta}{\epsilon}\right\} = 1\right\}. \tag{44}$$

By using the definition of $f(x,\zeta)$, i.e.,

$$f(x,\zeta) = \min\left\{\min_x \max\{b(x) - \zeta, 0\}, \min_x \chi_{\{x:b(x)<0\}}(x)\right\}, \tag{45}$$

where the characteristic function $\chi_R(x) = \infty$ if $x \neq R$ and 0, otherwise. Also, because of the fact that $\frac{\delta}{\epsilon} > 0$, we can arrive (44).■

In Theorem 3, it can be proved that set Z_R is a subset of he feasible region induced by a regular chance constraint, i.e., $Z_R \subseteq Z$ [38].

TABLE III
PARAMETERS SETTINGS

Parameters	Value
RLV Parameter a	1
RLV Parameter b	0
RLV Parameter ρ	0.5
Global accuracy index ψ	0.1
Global accuracy parameter ζ	10
Global iteration parameter κ	10
HD map parameter k	10
HD map size parameter r_1, r_2, r_3, and r_4	2, 3, 4, and 5 MB/m^2
Bandwidth B	$10MHz$
Transmission power w	23 dBm
Channel gain h	16
Gaussian Noise N_0	-96 dBm
Penalty parameters for four layers α_1, α_2, α_3, and α_4	4, 3, 2, and 1

VII. SIMULATION RESULTS

In this section, the simulation is performed for deterministic and uncertain capacity cases in Subsections VII-A and VII-B, respectively. The experiments are conducted on MATLAB platform. The RLV based utility function is defined as $u(t_i) = 2 \times (1 + t_i)^{0.5} - 1$, by setting $a = 0$, $b = 1$, and $\rho = 0.5$, as is suggested in [27]. For the federated analytics parameter settings, considering existing work [39], the constant parameter of iteration upper bound of federated analytics ζ is 20. The global accuracy index for federated analytics is 0.1. The parameters that denote HD map data size k is set as 10. The updating range for four layers are set as 2, 3, 4, and 5 MB/m^2, respectively. As for the communication model, as suggested in [3], the bandwidth B is 10 MHz. The transmission power is 23 dBm. The channel gain is 16. The Gaussian noise power is -96 dBm. And for the penalty function parameters, α_1, α_2, α_3, and α_4 are set as 1, 2, 3, and 4, for the highly dynamic layer, transient dynamic layer, transient layer, and permanent static layer, respectively. Overall, the general model parameter settings are summarized in Table III.

A. Deterministic Capacity Case

In this case, we firstly assume there are four vehicles in total, which will consume a random number of sensor data from the six types of sensors as introduced in Subsection III-B. We increase the edge server capacity, like available storage, from 100 MB to 1,000 MB by step 100 MB and the results are shown in Figs. 5 and 6, respectively. Fig. 5 illustrated the variation of allocation percentage of each HD map layer to the edge server. Obviously, the highly dynamic layer always has priority to consume the edge server resources and is the first one that achieves the 100% edge resources allocation at 300 MB capacity, followed by the transient

This article has been accepted for publication in IEEE Transactions on Intelligent Vehicles. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIV.2022.3224655

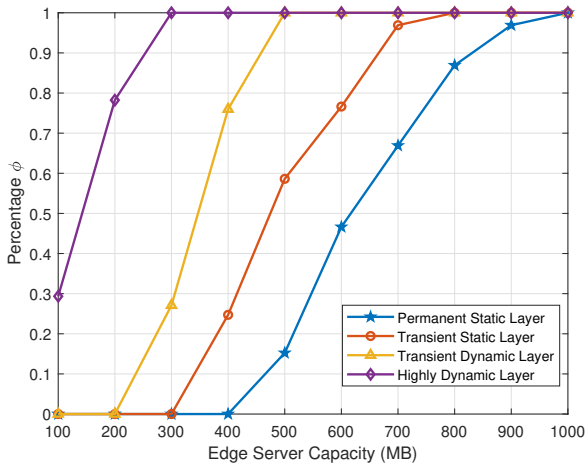


Fig. 5. Edge server allocation percentage with different edge server capacity.

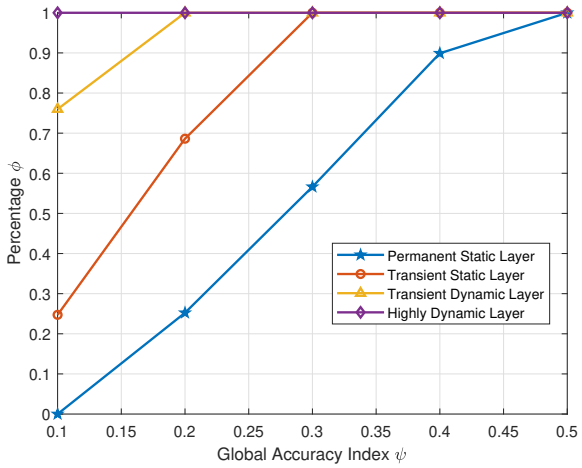


Fig. 7. Edge server allocation percentage with different global accuracy index.

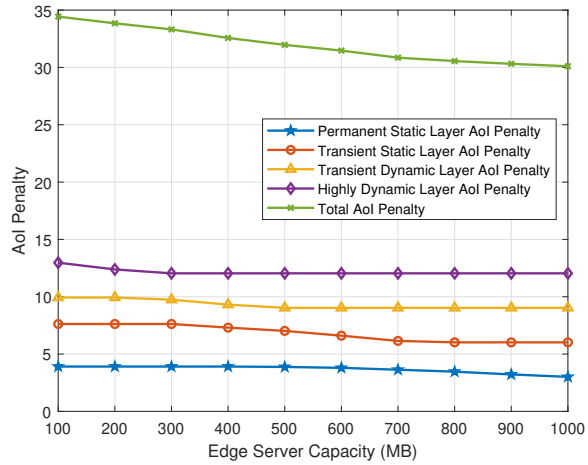


Fig. 6. The AoI penalty with different edge server capacity.

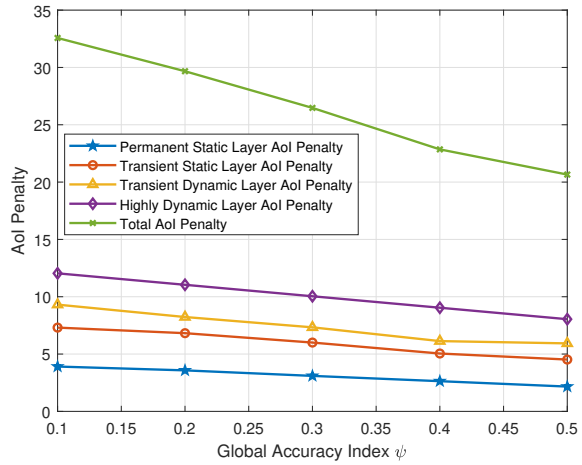


Fig. 8. The AoI penalty with different global accuracy index.

dynamic layer at 500 MB, the transient static layer at 800 MB, and the permanent static layer at 1000 MB, respectively. This is understandable because each layer has different latency requirements. For example, the highly dynamic layer, which consists of surrounding vehicles and pedestrians, is much more latency-sensitive than the permanent static layer which consists of road signs. Correspondingly, the penalty assigned to the highly dynamic layer will be more than the permanent static layer with regard to the same time delay. Typically, we set the penalty function to the highly dynamic layer, transient dynamic layer, transient static layer, and permanent static layer as quadruplicate, cubic, quadratic, and linear functions, respectively, which gives rise to the allocation priority order as in Fig. 5. In the beginning, when the edge server resources are limited, only the highly dynamic layer is able to consume the resources because small latency will result in the fourth power penalty. Thereafter, when the resources are plentiful, all of the four HD map layers can be transmitted via the edge server.

This change can also be demonstrated by Fig. 6, which shows the overall penalty and penalty of each layer. Generally, all of the curves indicate decreasing trend with the increase of edge server capacity. Because when the edge resources are more sufficient, more HD map data will be transmitted through the edge server such that the transmission delay can be reduced, resulting in the reduction of AoI penalty. In addition, we can see that the drop point of the highly dynamic layer AoI penalty is 100 MB edge capacity, 200 MB edge capacity for the transient dynamic layer, 300 MB edge capacity for the transient static layer, and 400 MB edge capacity for the permanent static layer, which also keeps consistent with the allocation priority order mentioned above.

Next, because the required HD map accuracy is crucial for federated analytics based HD map generation, we also explore the impact of the global accuracy index ψ on the performance. In this case, we keep the edge server capacity at 400 MB. The global accuracy index ψ increases from 0.1 to 0.5 by step 0.1, which represents local accuracy decreases gradually. The corresponding results are illustrated in Figs. 7 and 8. In Fig. 7, we can see that the overall trend is the percentage of

This article has been accepted for publication in IEEE Transactions on Intelligent Vehicles. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIV.2022.3224655

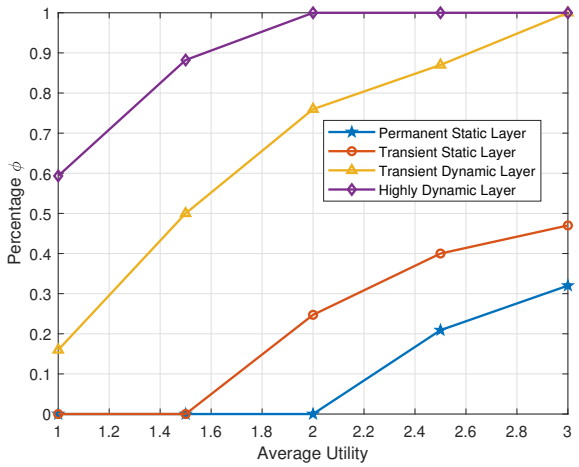


Fig. 9. Edge server allocation percentage with different average utility $\bar{U}$.

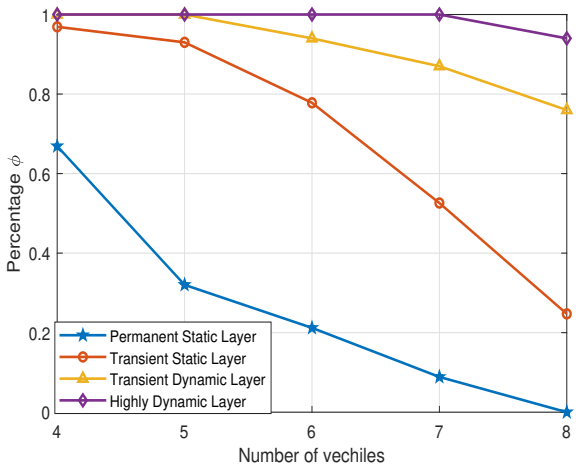


Fig. 11. Edge server allocation percentage with different number of vehicles.

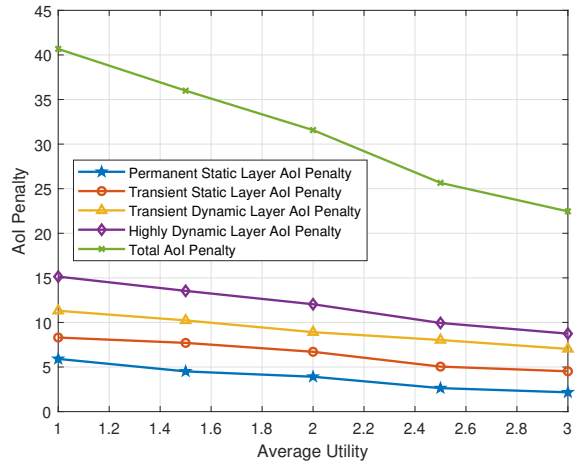


Fig. 10. The AoI penalty with different average utility $\bar{U}$.

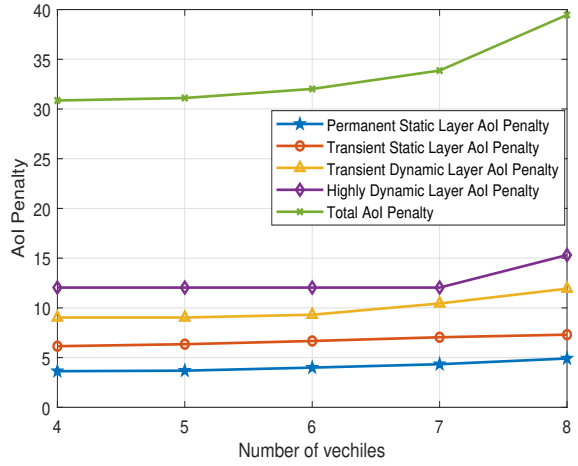


Fig. 12. The AoI penalty with different number of vehicles.

each layer allocated to edge server transmission increases with the increase of the global accuracy index. Since a high global accuracy index indicates a low local accuracy. Therefore, to achieve the required precision, the needed number of global aggregation will be less due to a lower accuracy desire. Accordingly, the generated data size will be reduced, resulting in a larger percentage of HD map data allocated to the edge server with a certain volume of capacity. Hence, with only 400 MB available capacity, all of the four layers can realize full edge server transmission with a lower accuracy expectation, i.e., ψ=0.5. In addition, similarly, because of the differences of the penalty functions, the allocation priority order still exists, where the highly dynamic layer has the highest priority, followed by the transient dynamic layer, transient static layer, and permanent layers. As for the results in Fig. 8, obviously, the AoI penalty decreases with the increase of global accuracy index for all cases. The reason is the same as the explanation for Fig. 7. The number of global aggregation will be reduced due to a lower accuracy requirement. As a consequence, the generated data size will be smaller, resulting in a lower transmission latency and AoI penalty. Also, this result reveals that a trade-off can be made between accuracy and transmission delay so the decisions can be adjusted for cases with specific requirements.

Intuitively, HD map generation heavily relies on crowdsourcing vehicle sensor data. Therefore, we also discuss the influence of different average utilities among HD map allocation and the AoI penalty. In this part, we keep the available edge server capacity as 400 MB and the global accuracy index as 0.1. The average utility is increased from 1 to 3 by step 0.5. The results are shown in Figs. 9 and 10. In Fig. 10, we can see the allocation percentage increases with the increase in average utility. This is because when the average utility is larger, the needed number of federated analytics aggregation will be less regarding a given accuracy level. Thus, the data size that needs to be transmitted will be less as well, resulting in lower latency and AoI penalty. Likewise, in any case, the highly dynamic layer is always allocated the majority of edge resources, which brings into correspondence with previous experiments. Also, the penalty is illustrated in Fig. 10. The AoI penalty decreases

TABLE IV
RESULTS FOR UNCERTAIN CAPACITY CASE

Parameter		Optimal Solution		CVaR Model			Random		
ϵ	δ	Opt.Val	Time (s)	Value	Gap	Time (s)	Value	Gap	Time (s)
0.05	0.01	32.57	4.73	32.93	1.11%	0.05	36.49	12.04%	0.02
0.05	0.02	32.57	2.32	33.02	1.38%	0.06	33.69	3.44%	0.02
0.1	0.01	32.57	9.16	32.97	1.23%	0.06	36.03	10.62%	0.02
0.1	0.02	32.57	10.85	33.68	3.17%	0.05	34.87	7.06%	0.02

with the increase of average utility $\bar{U}$. Since a higher average utility leads to smaller data size, the reduction in transmission time will bring less AoI penalty.

In addition, we investigate on the impact of number of vehicles. We set the edge server capacity as 700 MB. The global accuracy index ψ is o.1. We increase the number of vehicles from 4 to 8 by step 1. Fig. 11 shows the percentage allocation with the increase of number of vehicles. All the curves shows a decreasing tendency. This is because when the more and more vehicles try to connect with edge server for uploading data, the limited capacity cannot meet all of the layers latency requirements of all the vehicles. But the highly dynamic layer always has the priority to be offloaded to the edge server, even when the number of vehicle achieves 7, the highly dynamic layer will be transmitted 100 percent, which is also in accordance with previous findings. Fig. 12 illustrates the tendency of the AoI penalty with the increase of vehicles. All of the curves show an increasing trend because the limited capacity cannot handle all the layers transmission when the vehicle number is sufficiently large. Besides, another thing is, regarding the amplitude of change, the highly dynamic layer is the most significant one. Since the highly dynamic layer is the most latency-sensitive one compared with the other layers. When the latency increases, the corresponding penalty will increase exponentially.

B. Uncertain Capacity Case

In this subsection, we will use the CVaR based inner approximation to estimate the uncertain capacity case. The mean value of edge server capacity is set as 400 MB. The risk parameter ϵ is chosen from $\{0.05, 0.1\}$. And the Wasserstein distance δ is set as $\{0.1, 0.02\}$. The number of instances N is 10. In addition CVaR based inner approximation, we also use random strategy to act as a comparison method. Also, since we can refer to the optimal solution from the deterministic case, the optimality gap can be adopted as the evaluation metric, which is defined as GAP = $\frac{|\text{Value - Opt.Val}|}{\text{Opt.Val}}$, where Value indicates the inner approximated AoI penalty and the Opt.Val is the optimal AoI penalty based on the deterministic case. The results are shown in Table IV.

Generally, we can see that no matter for CVaR model based approximation or random strategy, they have a certain amount of accuracy gap regarding optimal solution. The CVaR model is usually $1\% - 3\%$ away from the optimality. However, the optimality gap for random strategy varies a lot, where the least one is 3.44 % and the largest one is 12.04%, because the allocation percentage is randomly determined. Therefore, in terms of approximation accuracy, CVaR model is better. Since the existence of uncertainty, the optimality gap is inevitable but $1\% - 3\%$ difference is also acceptable. In addition, the results for CVaR model shows that the least gap is achieved when $\epsilon = 0.05$ and $\delta = 0.01$ while the largest gap comes from the case with $\epsilon = 0.1$ and $\delta = 0.02$. This is also understandable because the larger value indicates a more relaxed or larger ambiguity space. If the ϵ is smaller, the risk tolerance is smaller as well, so high risk instances will be exempted. While for δ, if the number is larger, the radius of the uncertain Wasserstein ball is larger, which increases the ambiguity level.

Another thing that needs to be noted is the execution time. To obtain the optimal solution, the time consumption can be up to 10 seconds. However, the CVaR model execution time is typically less than one second. The reason behind this might be CVaR model is a second-order conic programming and does not involve any binary variables. Therefore, for time-sensitive applications, the CVaR model based inner approximation can be an effective approach for the trade-off between time consumption and accuracy. Additionally, random strategy achieves the lowest execution time because it does not solve any complex programming but randomly chooses a percentage within the given range. Overall, the results in Table IV demonstrate that the CVaR model based approximation is able to find near-optimal solutions with much less time consumption.

VIII. CONCLUSION

As a latency-sensitive application, HD map transmission needs to be performed in a timely manner. This paper investigates an information staleness minimization problem for federated analytics based HD map generation and layer-wise transmission offloading. Fortunately, emerging MEC provides a low-latency paradigm for data transmission. But the available edge computing resources may be variational practically. Therefore, this paper discusses two cases, i.e., the deterministic edge capacity case and the uncertain edge capacity case. For the deterministic edge capacity case, the optimal solution is obtained analytically. The influence of different edge server capacity, federated analytics accuracy, vehicle utility, and number of vehicles are also discussed. For the uncertain edge capacity case, the problem is reformulated into a DRCCO optimization problem and solved by the CVaR model based approximation. The experiments demonstrate the CVaR model based approximation is able to find near-optimal solutions with much less time consumption.

REFERENCES

[1] National Highway Traffic Safety Administration (NHTSA). Automated vehicles for safety. [Online]. Available: https:

//www.nhtsa.gov/technology-innovation/automated-vehicles-safety\#issue-road-self-driving.

[2] R. Liu, J. Wang, and B. Zhang, "High definition map for automated driving: Overview and analysis," *The Journal of Navigation*, vol. 73, no. 2, pp. 324–341, March 2020.

[3] D. Chen, D. Wang, Y. Zhu, and Z. Han, "Digital twin for federated analytics using a bayesian approach," *IEEE Internet of Things Journal*, vol. 8, no. 22, pp. 16 301–16 312, July 2021.

[4] F. Jomrich, J. Schmid, S. Knapp, A. Höß, R. Steinmetz, and B. Schuller, "Analysing communication requirements for crowd sourced backend generation of hd maps used in automated driving," in *IEEE Vehicular Networking Conference (VNC)*, Taipei, Taiwan, China, December 2018.

[5] S. Kokovin, J.-F. Thisse, and E. Zhelobodko, "Monopolistic competition: Beyond the ces," *CEPR Discussion Paper No. DP7947*, August 2010.

[6] X. Xu, S. Gao, and M. Tao, "Distributed online caching for high-definition maps in autonomous driving systems," *IEEE Wireless Communications Letters*, vol. 10, no. 7, pp. 1390–1394, July 2021.

[7] F. Zhang, J. Zhai, X. Shen, O. Mutlu, and X. Du, "Poclib: A high-performance framework for enabling near orthogonal processing on compression," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 2, pp. 459–475, June 2021.

[8] D. Chen, C. S. Hong, L. Wang, Y. Zha, Y. Zhang, X. Liu, and Z. Han, "Matching-theory-based low-latency scheme for multitask federated learning in mec networks," *IEEE Internet of Things Journal*, vol. 8, no. 14, pp. 11 415–11 426, July 2021.

[9] G. Lloret-Talavera, M. Jorda, H. Servat, F. Boemer, C. Chauhan, S. Tomishima, N. N. Shah, and A. J. Pena, "Enabling homomorphically encrypted inference for large dnn models," *IEEE Transactions on Computers*, vol. 71, no. 5, pp. 1145–1155, August 2021.

[10] H. Chu, L. Guo, B. Gao, H. Chen, N. Bian, and J. Zhou, "Predictive cruise control using high-definition map and real vehicle implementation," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 11 377–11 389, September 2018.

[11] M. Hirabayashi, A. Sujiwo, A. Monrroy, S. Kato, and M. Edahiro, "Traffic light recognition using high-definition map features," *Robotics and Autonomous Systems*, vol. 111, pp. 62–72, January 2019.

[12] Z. Jian, S. Zhang, S. Chen, X. Lv, and N. Zheng, "High-definition map combined local motion planning and obstacle avoidance for autonomous driving," in *IEEE Intelligent Vehicles Symposium (IV)*, Paris, France, June 2019, pp. 2180–2186.

[13] Y. Yoon, H. Chae, and K. Yi, "High-definition map based motion planning, and control for urban autonomous driving," *SAE Technical Paper*, 2021.

[14] F. Wang, D. Guan, L. Zhao, and K. Zheng, "Cooperative v2x for high definition map transmission based on vehicle mobility," in *IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*, Kuala Lumpur, Malaysia, April 2019.

[15] X. Xu, S. Gao, and M. Tao, "Distributed online caching for high-definition maps in autonomous driving systems," *IEEE Wireless Communications Letters*, vol. 10, no. 7, pp. 1390–1394, July 2021.

[16] Y. Wu, Y. Shi, Z. Li, and S. Chen, "A cluster-based data offloading strategy for high definition map application," in *IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, May 2020.

[17] Y. Liu, M. Song, and Y. Guo, "An incremental fusing method for high-definition map updating," in *IEEE International Conference on Systems, Man and Cybernetics (SMC)*, Bari, Italy, October 2019, pp. 4251–4256.

[18] J. Xie, J. Tang, and S. Liu, "An energy-efficient high definition map data distribution mechanism for autonomous driving," *arXiv preprint arXiv:2010.05233*, 2020.

[19] J. Jiao, "Machine learning assisted high-definition map creation," in *IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, vol. 1, Yokyo, Japan, July 2018, pp. 367–373.

[20] C. Kim, S. Cho, M. Sunwoo, P. Resende, B. Bradaï, and K. Jo, "Updating point cloud layer of high definition (hd) map based on crowd-sourcing of multiple vehicles installed lidar," *IEEE Access*, vol. 9, pp. 8028–8046, 2021.

[21] H. Zhuang, C. Wang, Y. Qian, and M. Yang, "Semantic-based road segmentation for high-definition map construction," in *International Conference on Cognitive Systems and Signal Processing*, Zhuhai, China, December 2020, pp. 527–537.

[22] Y. Li, C. Wang, F. Chen, Z. Su, and X. Wu, "A method for generating large-scale high definition color-point map," in *IEEE International Conference on Real-time Computing and Robotics (RCAR)*, Irkutsk, Russia, August 2019, pp. 487–492.

[23] Automotive Edge Computing Consortium (AECC) White Paper. (2020, May) Operational behavior of a high definition map application. [Online]. Available: https://aecc.org/wp-content/uploads/2020/07/Operational_Behavior_of_a_High_Definition_\\Map_Application.pdf

[24] Y. B. Can, A. Liniger, O. Unal, D. Paudel, and L. Van Gool, "Understanding bird's-eye view semantic hd-maps using an onboard monocular camera," *arXiv preprint arXiv:2012.03040*, 2020.

[25] W. Feng, J.-M. Friedt, G. Cherniak, and M. Sato, "Batch compressive sensing for passive radar range-doppler map generation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 6, pp. 3090–3102, February 2019.

[26] F. Ghallabi, G. El-Haj-Shhade, M.-A. Mittet, and F. Nashashibi, "Lidar-based road signs detection for vehicle localization in an hd map," in *IEEE Intelligent Vehicles Symposium (IV)*, Paris, France, June 2019, pp. 1484–1490.

[27] Y. Zhao, H. Wang, H. Su, L. Zhang, R. Zhang, D. Wang, and K. Xu, "Understand love of variety in wireless data market under sponsored data plans," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 4, pp. 766–781, February 2020.

[28] K. Yang, T. Jiang, Y. Shi, and Z. Ding, "Federated learning via over-the-air computation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 2022–2035, January 2020.

[29] C. Ma, J. Konečnỳ, M. Jaggi, V. Smith, M. I. Jordan, P. Richtárik, and M. Takáč, "Distributed optimization with arbitrary local solvers," *Optimization Methods and Software*, vol. 32, no. 4, pp. 813–848, February 2017.

[30] R. Han, M. Si, J. Demmel, and Y. You, "Dynamic scaling for low-precision learning," in *Proceedings of the 26th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*, Virtual Event Republic of Korea, February 2021, pp. 480–482.

[31] N. H. Tran, W. Bao, A. Zomaya, M. N. Nguyen, and C. S. Hong, "Federated learning over wireless networks: Optimization model design and analysis," in *IEEE Conference on Computer Communications (INFOCOM)*, Paris, France, April 2019, pp. 1387–1395.

[32] S. Zhang, H. Zhang, Z. Han, H. V. Poor, and L. Song, "Age of information in a cellular internet of uavs: Sensing and communication trade-off design," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6578–6592, October 2020.

[33] F. Wu, H. Zhang, J. Wu, Z. Han, H. V. Poor, and L. Song, "Uav-to-device underlay communications: Age of information minimization by multi-agent deep reinforcement learning," *IEEE Transactions on Communications*, vol. 69, no. 7, pp. 4461–4475, July 2021.

[34] Z. Chen, D. Kuhn, and W. Wiesemann, "Data-driven chance constrained programs over wasserstein balls," *arXiv preprint arXiv:1809.00210*, 2018.

[35] F. Luo and S. Mehrotra, "Decomposition algorithm for distributionally robust optimization using wasserstein metric with an application to a class of regression models," *European Journal of Operational Research*, vol. 278, no. 1, pp. 20–35, October 2019.

[36] J. Blanchet and K. Murthy, "Quantifying distributional model risk via optimal transport," *Mathematics of Operations Research*, vol. 44, no. 2, pp. 565–600, May 2019.

[37] W. Xie and S. Ahmed, "Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1860–1867, July 2017.

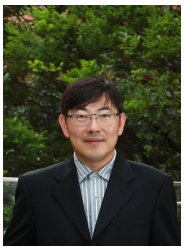
[38] W. Xie, "On distributionally robust chance constrained programs with wasserstein distance," *Mathematical Programming*, vol. 186, pp. 115–155, November 2019.

[39] Y. Yu, D. Chen, X. Tang, T. Song, C. S. Hong, and Z. Han, "Incentive framework for cross-device federated learning and analytics with multiple tasks based on a multi-leader-follower game," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 5, pp. 3749–3761, July 2022.

This article has been accepted for publication in IEEE Transactions on Intelligent Vehicles. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIV.2022.3224655

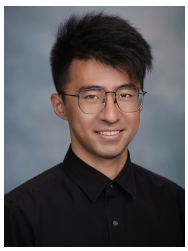
Dawei Chen (S'19-M'22) received the B.S. degree from the Huazhong University of Science and Technology, Wuhan, China, and the Ph.D. degree from the University of Houston, Houston, TX, USA, in 2015 and 2021, respectively. From 2015 to 2016, he was a network engineer at Ericsson. Currently he is a research scientist at Toyota Motor North America, InfoTech Labs. His research interests include deep learning, federated learning/analytics cloud/edge computing, and wireless networks.

Yifei Zhu is currently an Assistant Professor at the University of Michigan-Shanghai Jiao Tong University Joint Institute in Shanghai Jiao Tong University, China. He received the B.E. degree from Xi'an Jiaotong University, China, in 2012, and the M.Phil. degree from The Hong Kong University of Science and Technology, China, in 2015, and the Ph.D degree in Computer Science from the Simon Fraser University, Canada, in 2020. His current research interests include edge computing, multimedia networking, and distributed machine learning systems, where he published in ACM SIGCOMM, IEEE INFOCOM, ACM Multimedia, and many other venues.



Dan Wang (S'05–M'07–SM'13) received the BSc degree from Peking University, Beijing, China, in 2000, MSc degree from Case Western Reserve University, Cleveland, Ohio, in 2004, and PhD degree from Simon Fraser University, Burnaby, British Columbia, Canada, in 2007, all in computer science.

Dan Wang's research falls in general computer networking and systems, where he published in ACM SIGCOMM, ACM SIGMETRICS and IEEE INFOCOM, and many others. He is the steering committee chair of IEEE/ACM IWQoS. He served as the TPC co-Chair of IEEE/ACM IWQoS 2020. His recent research focus on smart energy systems. He won the Best Paper Awards of ACM e-Energy 2018 and ACM Buildsys 2018. He has served as a TPC co-Chair of the ACM e-Energy 2020 and he will serve as General co-Chair of the ACM e-Energy 2022. He is a steering committee member of ACM e-Energy. He serves as a founding area editor of ACM SIGEnergy Energy Informatics Review. His research has been adopted by industry, e.g., Henderson, Huawei, and IBM. He won the Global Innovation Award, TechConnect, in 2017.



Haoxin Wang (S'16-M'21) is an Assistant Professor in the Department of Computer Science at Georgia State University, and leads the Advanced Mobility & Augmented Intelligence (AMAI) Lab. He received the Ph.D. degree from The University of North Carolina at Charlotte in 2020. From 2020 to 2022, he was a Research Scientist at Toyota Motor North America, InfoTech Labs. His current research interests include mobile AR/VR, autonomous driving, holographic communications, and edge intelligence.

Jiang Xie (Fellow, IEEE) received the B.E. degree from Tsinghua University, Beijing, China, the M.Phil. degree from the Hong Kong University of Science and Technology, and the M.S. and Ph.D. degrees from the Georgia Institute of Technology, all in electrical and computer engineering. She joined the Department of Electrical and Computer Engineering, the University of North Carolina at Charlotte (UNC Charlotte) as an Assistant Professor in August 2004, where she is currently a Full Professor. Her current research interests include resource and mobility management in wireless networks, mobile computing, Internet of Things, cloud/edge computing, and virtual/augmented reality. She received the U.S. National Science Foundation NSF Faculty Early Career Development (CAREER) Award in 2010, the Best Paper Award from IEEE Global Communications Conference in 2017, the Best Paper Award from IEEE/WIC/ACM International Conference on Intelligent Agent Technology in 2010, and the Graduate Teaching Excellence Award from the College of Engineering at UNC-Charlotte in 2007. She is on the editorial boards of the IEEE Transactions on Wireless Communications, IEEE Transactions on Sustainable Computing and Journal of Network and Computer Applications (Elsevier). She is a Senior Member of ACM.

Xiao-ping Zhang (Fellow, IEEE) received B.S. and Ph.D. degrees from Tsinghua University, in 1992 and 1996, respectively, both in Electronic Engineering. He holds an MBA in Finance, Economics and Entrepreneurship with Honors from the University of Chicago Booth School of Business, Chicago, IL.

Since Fall 2000, he has been with the Department of Electrical, Computer and Biomedical Engineering, Toronto Metropolitan University (Formerly Ryerson University), Toronto, ON, Canada, where he is currently a Professor and the Director of the Communication and Signal Processing Applications Laboratory. He has served as the Program Director of Graduate Studies. He is cross-appointed to the Finance Department at the Ted Rogers School of Management, Ryerson University. He was a Visiting Scientist with the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA, USA, in 2015 and 2017. He is a frequent consultant for biotech companies and investment firms. His research interests include image and multimedia content analysis, sensor networks and IoT, machine learning, statistical signal processing, and applications in big data, finance, and marketing.

Dr. Zhang is Fellow of the Canadian Academy of Engineering, Fellow of the Engineering Institute of Canada, Fellow of the IEEE, a registered Professional Engineer in Ontario, Canada, and a member of Beta Gamma Sigma Honor Society. He is the general Co-Chair for the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2021. He is the general co-chair for 2017 GlobalSIP Symposium on Signal and Information Processing for Finance and Business, and the general co-chair for 2019 GlobalSIP Symposium on Signal, Information Processing and AI for Finance and Business. He was an elected Member of the ICME steering committee. He is the General Chair for the IEEE International Workshop on Multimedia Signal Processing, 2015. He is Editor-in-Chief for the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING. He is Senior Area Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING. He served as Senior Area Editor the IEEE TRANSACTIONS ON SIGNAL PROCESSING and Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the IEEE SIGNAL PROCESSING LETTERS. He is selected as IEEE Distinguished Lecturer by the IEEE Signal Processing Society and by the IEEE Circuits and Systems Society.

Zhu Han (S'01–M'04-SM'09-F'14) received the B.S. degree in electronic engineering from Tsinghua University, in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, in 1999 and 2003, respectively.

From 2000 to 2002, he was an R&D Engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a Research Associate at the University of Maryland. From 2006 to 2008, he was an assistant professor at Boise State University, Idaho. Currently, he is a John and Rebecca Moores Professor in the Electrical and Computer Engineering Department as well as in the Computer Science Department at the University of Houston, Texas. Dr. Han's main research targets on the novel game-theory related concepts critical to enabling efficient and distributive use of wireless networks with limited resources. His other research interests include wireless resource allocation and management, wireless communications and networking, quantum computing, data science, smart grid, security and privacy. Dr. Han received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, IEEE Leonard G. Abraham Prize in the field of Communications Systems (best paper award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. Dr. Han was an IEEE Communications Society Distinguished Lecturer from 2015-2018, AAAS fellow since 2019, and ACM distinguished Member since 2019. Dr. Han is a 1% highly cited researcher since 2017 according to Web of Science. Dr. Han is also the winner of the 2021 IEEE Kiyo Tomiyasu Award, for outstanding early to mid-career contributions to technologies holding the promise of innovative applications, with the following citation: "for contributions to game theory and distributed management of autonomous communication networks."