Federated Learning-Based Joint Radar-Communication mmWave Beamtracking with Imperfect CSI for V2X Communications

1st Sanjay Bhardwaj
ICT Convergence Research Center
Department of IT Convergence Engineering
Kumoh National Institute of Technology
Gumi, South Korea
sanjay.b1976@gmail.com

2nd Dong-Seong Kim
ICT Convergence Research Center
Department of IT Convergence Engineering
Kumoh National Institute of Technology
Gumi, South Korea
dskim@kumoh.ac.kr

***Abstract*—Robust beamforming is used to ensure communication quality, particularly for vehicle-to-everything (V2X) communications. Moreover, millimeter waves (mmWaves) are used in a new era of communication, providing a high data rate and high spectrum efficiency. However, mmWave links are prone to short-range communication because of serious free-space path loss and blockage by obstacles. In addition, owing to constraints on bandwidth and privacy considerations, it is computationally difficult to transmit the entire dataset between vehicle users and road side units (RSUs). As a result, FL-BT, a federated learning (FL)-based joint radar and communication mmWave beamtracking with imperfect channel state information (CSI) for V2X communications, is proposed. FL-BT is a federated module at vehicle user (local clients) and an RSU (main server) consisting of support vector regression (SVR) along with an unscented Kalman filter (UKF) that predicts the distribution of future channel behavior based on a sequence of input signals available. The simulation results demonstrate that the proposed FL-BT approach offers significant improvements in V2X communication performance by incorporating the UKF and SVR in a state-space model for precise future channel-value prediction and a federated learning framework for handling bandwidth and privacy considerations. This leads to a more robust and accurate beamtracking performance compared to traditional beamtracking methods.**

***Index Terms*—Federated Learning, Imperfect channel state estimation, Joint Radar-Communication, Millimeter wave, Support Vector Regression, Unscented Kalman Filter,**

I. Introduction

A novel networked machine learning (ML) paradigm termed federated learning (FL), which respects user privacy while making use of user device data and computational power. FL has been demonstrated to be a useful technique for helping millimeter wave (mmWave) communication system establish a link successfully [1]. MmWave technology is attractive because it can support large data rates in wireless communication networks, such as vehicle-to-everything (V2X) connections. Modern mmWave communication systems can use machine-learning techniques to analyze large amounts of wireless traffic data and learn how to adapt to changing conditions, which is becoming increasingly important with the rise of more wireless devices [2].

This work was supported by Priority Research Centers Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (MEST) (2018R1A6A1A03024003) and the MSIT (Ministry of Science and ICT), Korea, under the Grand Information Technology Research Center support program (IITP-2022-2020-0-01612) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation.

II. Related Work

Next-gen applications like V2X connectivity rely on both radar and mmWave communication. MmWave auto radars offer high-res sensing with a wide view, while mmWave comms provide high-data-rate solutions for connected cars [3], [4]. JRC (joint radar-communication) technology merges mmWave automotive radars and mmWave communications, resulting in reusable hardware and similar signaling waveform. This leads to improved power usage, spectrum efficiency, and market penetration [5]. The authors of [6] proposed ML-BT, a machine-learning-based beamtracking technique for mmWave vehicle transmissions that uses long short-term memory (LSTM) networks. However, the method has limitations, such as longer training times, high memory requirements, overfitting risk, and difficulties in implementing Dropout, which results in less than optimal outcomes [7], [8].

The estimation of the AoA in a mobile mmWave channel was proposed by [9] using an extended Kalman filter (EKF)-based technique known as EKF-BT. The EKF linearizes the equations by approximating the non-linear functions with their first-order Taylor series expansions around the current estimate of the state. This approach can lead to inaccuracies in the linearization, particularly if the state estimates are far from the true state of the system [10]. In contrast to those from the EKF method, the state estimates from the UKF technique reflect the state uncertainty more accurately UKF because it uses a set of sigma points to represent the probability distribution of the system state. These sigma points are chosen such that they capture the mean and covariance of the state distribution and are propagated through the system dynamics to estimate the predicted state. The measurements are then used to correct the

predicted state, and a weighting factor is applied to each sigma point to obtain the final estimate of the state [11].

The UKF has two key advantages. Firstly, it can handle nonlinear systems, unlike the traditional Kalman filter which linearizes the system dynamics around the mean of the state estimate. The UKF propagates the sigma points through the nonlinear system model to obtain the predicted state, resulting in accurate estimation for nonlinear systems [12]. Secondly, the UKF can handle systems with higher-dimensional state spaces, which is difficult for the traditional Kalman filter that requires a linear relationship between the state and measurement for estimation [13]. The UKF is widely used in navigation, target tracking, guidance, control, and signal processing. It is a popular technique in robotics, automated vehicles, and aerospace applications, where non-linearity and high-dimensional state spaces are common [14].

Support vector regression (SVR) effectively handles model overfitting and poor generalization, resulting in strong generalization capability and a global optimal solution [15]. It has gained popularity for predicting time series due to its ability to handle non-linearity and high-dimensional data [16]. SVR models underlying patterns and trends in time series data to make predictions about future values. It finds the best linear regression function that can approximate the data and can handle both linear and non-linear time series data [17]. Advantages of using SVR for time series prediction include its ability to handle non-linearity and high dimensional data, and robustness to noisy or outlier data [18].

As a result, the UKF is used to perform recursive state estimates on the nonlinear state-space model based on SVR with significant stochastic uncertainty. The enhanced beamtracking method utilizes federated learning to model rapidly varying mmWave channels, and it uses data from a vehicle user to illustrate the continuous evolution of the AoA and AoD. In particular, using the order of previous (past) channel estimations, the federated module (composed of SVR and UKF) forecasts the distribution of the AoA and AoD states. Our findings show that the proposed method outperforms traditional beamtracking techniques under a variety of high-mobility and imperfect CSI conditions. The rest of the contributions to this paper are summarized as follows;

- To improve the beamtracking performance of mmWave systems for V2X communications, our technique combines SVR and UKF as a federated module. The federated module can appropriate the complicated dynamic channel behavior brought on by a vehicle's user mobility as opposed to commonly utilized basic linear models, which results in improved beamtracking performance. Numerical analysis was used to confirm that the proposed beamtracking approach is better than current methods.
- To study channel dynamics and forecast future channel states, this study's fundamental idea is to use a federated learning architecture. We included the SVR-based prediction model, which was further improved by the UKF framework. In the measurement update phase, we used a mathematical model that defines the relationship between the transmitted beam and the received signal. It should be emphasized that machine learning models are only able to reproduce the dynamic behavior of channels.
- A comparison was made with ML-BT [6], EKF-BT was proposed by [9] and with an approach that has perfect CSI, called FL-BT-P-CSI.

The rest of the paper is organized as: Section III presents Federated communication model, Performance evaluation is explained in Section IV and finally, conclusion is presented in Section V.

III. Federated Communication Model

This section elaborates on the proposed system model, along with the radar communication model, and expands on the proposed approach by explaining the SVR, UKF, and FL.

A. System and Radar Communication Model

In a V2X scenario, Fig. 1 considers a mmWave MIMO communication system (communication module) assisted by a MIMO radar system (radar module). SVR and UKF as a federated learning architecture (federated module) are also displayed beside it at each car user (local clients) and at RSU (main server). Considering RSU and vehicle users (VUs) with antenna array size of N_r and N_v, respectively with N_{RF} RF chains. The mmWave channel model for downlink from RSU to VUs with a channel matrix of H for size $N_r \times N_v$, [19]

$$\mathbf{H} = \sqrt{\frac{N_r N_v}{P}} \sum_{p=1}^{P} \psi_p \mathbf{a}^{(v)} \left(\phi_p^{(v)}\right) \left(\mathbf{a}^{(r)} \left(\phi_p^{(r)}\right)\right)^H, \quad (1)$$

where P represents the total number of paths, ψ_p represents the complex channel gain for p^{th} path, $\left(\phi_p^{(v)}\right)$ represents AoA, and $\left(\phi_p^{(r)}\right)$ represents AoD. The steering vectors $\mathbf{a}^{(r)}\left(\phi^{(r)}\right)$ and $\mathbf{a}^{(v)}\left(\phi^{(v)}\right)$ are given as $\mathbf{a}^{(r)}\left(\phi^{(r)}\right) = \frac{1}{\sqrt{N_r}}\left[1, e^{\frac{j2\pi d_r cos\phi}{\lambda}}, e^{\frac{j2\pi 2d_r cos\phi}{\lambda}}, ...e^{\frac{j2\pi(N_r-1)d_r cos\phi}{\lambda}}\right]^T$, and $\mathbf{a}^{(v)}\left(\phi^{(v)}\right) = \frac{1}{\sqrt{N_v}}\left[1, e^{\frac{j2\pi d_v cos\phi}{\lambda}}, e^{\frac{j2\pi 2d_v cos\phi}{\lambda}}, ...e^{\frac{j2\pi(N_v-1)d_v cos\phi}{\lambda}}\right]^T$, where λ is the wavelength of the signal, d_r and d_v is the distance between adjacent antennas at RSU and VU, respectively, given as $\frac{\lambda}{2}$. Eq. (1) in the matrix form can be written as

$$\mathbf{H} = \mathbf{A}_R \mathbf{H}_a \mathbf{A}_V^H, \quad (2)$$

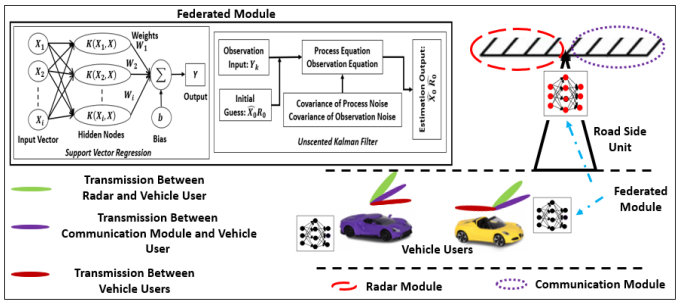

Fig. 1: Elucidation of System Model

where $\mathbf{H}_a = \sqrt{\frac{N_r N_v}{P}} diag\left(\psi_1, \psi_2, ..., \psi_P\right)$, $\mathbf{A}_R = \left[\mathbf{a}^r\left(\phi_1^r\right), \mathbf{a}^r\left(\phi_2^r, ...\mathbf{a}^r\left(\phi_P^r\right)\right)\right] \in \mathbb{C}^{N_r \times P}$, $\mathbf{A}_V = \left[\mathbf{a}^v\left(\phi_1^v\right), \mathbf{a}^v\left(\phi_2^v, ...\mathbf{a}^v\left(\phi_P^v\right)\right)\right] \in \mathbb{C}^{N_v \times P}$. In order to investigate the effect of errors in CSI (imperfect CSI), the CSI $\hat{\mathbf{H}}$ is subjected to noise errors. Therefore, the estimated channel based on the imperfect CSI is represented as [20], $\hat{\mathbf{H}} = \xi\mathbf{H} + \sqrt{1-\xi^2}\mathbf{E}$ which results in $\mathbf{H} = \frac{\hat{\mathbf{H}} - \sqrt{1-\xi^2}\mathbf{E}}{\xi}$, where $\xi \in [0,1]$ is the CSI accuracy, higher value of ξ denotes higher accuracy and $\mathbf{E}$ is the channel estimation matrix whose entries follow i.i.d. $\mathscr{CN}(0,1)$.

At the b^{th} beam transmission, RSU apply N_r^B different pilot beam patterns given as $\mathbf{f}_b \in \mathbb{C}^{N_r \times 1}$ for $b = 1, 2, ...N_r^B$ and VU apply N_v^B different pilot beam patterns given as $\mathbf{w}_c \in \mathbb{C}^{N_v \times 1}$ for $c = 1, 2, ...N_v^B$, therefore the total number of slots required by the received beam patterns to scan over is $N_v^D = \frac{N_v^B}{N_{RF}}$. Thus, the received vector signal for the b^{th} pilot beam in the s^{th} time slot is given as

$$\mathbf{y}_{s,b} = \mathbf{W}_s^H \mathbf{H}\mathbf{f}_b x_{tr} + \mathbf{W}_s^H \mathbf{n}_{s,b}, \tag{3}$$

where $\mathbf{y}_{s,b} \in \mathbb{C}^{N_{RF} \times 1}$ for $s \in \{1, 2, ...N_v^D\}$, $\mathbf{W}_s = \left[\mathbf{w}_{(s-1)N_{RF}+1}, ...\mathbf{w}_{sN_{RF}}\right] \in \mathbb{C}^{N_v \times N_{RF}}$ is the receive beam pattern matrix for $\mathbf{f}_b$, x_{tr} is the transmitted pilot symbol, $\mathbf{H} \in \mathbb{C}^{N_r \times N_v}$ is the channel matrix and $\mathbf{n}_{s,b} \in \mathbb{C}^{N_v \times 1}$ is the noise vector.

Combination of the signal vectors $\mathbf{y}_{s,b}$ over entire slots $s \in \{1, 2, ...N_v^D\}$, the complete signal $\mathbf{y}_b \in \mathbb{C}^{N_v^B \times 1}$ for $\mathbf{f}_b$ pilot beam pattern is given as

$$\mathbf{y}_b = \mathbf{W}^H \mathbf{H}\mathbf{f}_b x_p + diag\left(\mathbf{W}_1^H, \mathbf{W}_2^H, ...\mathbf{W}_{N_v^D}^H\right) \times \left[\mathbf{n}_{1,b}^T, \mathbf{n}_{2,b}^T, ...\mathbf{n}_{N_v^D,b}^T\right]^T, \tag{4}$$

where $\mathbf{W} = \left[\mathbf{W}_1, \mathbf{W}_2, ...\mathbf{W}_{N_v^D}\right] \in \mathbb{C}^{N_v \times N_v^D}$. Therefore, for all N_r^B transmit beams, the collection of $\mathbf{y}_b, b \in \{1, 2, ...N_r^B\}$ gives

$$\mathbf{Y} = \mathbf{W}^H \mathbf{H}\mathbf{F}\mathbf{X} + \mathbf{N} = \sqrt{Q}\mathbf{W}^H \mathbf{H}\mathbf{F} + \mathbf{N}, \tag{5}$$

where $\mathbf{X} = \sqrt{Q}\mathbf{I}_{N_r^B} \in \mathbb{C}^{N_r^B \times N_r^B}$ for Q is the pilot power, $\mathbf{Y} = \left[\mathbf{y}_1, \mathbf{y}_2, ...\mathbf{y}_{N_r^B}\right] \in \mathbb{C}^{N_r^B \times N_v^B}, \mathbf{F} = \left[\mathbf{f}_1, \mathbf{f}_2, ...\mathbf{f}_{N_r^B}\right] \in \mathbb{C}^{N_r \times N_r^B} \mathbf{N} \in \mathbb{C}^{N_r^B \times N_v^B}, \mathbf{N} = diag\left(\mathbf{W}_1^H, \mathbf{W}_1^H, ...\mathbf{W}_{N_r^B}^H\right) \left[\left[\mathbf{n}_{1,1}^T, \mathbf{n}_{2,1}^T, ...\mathbf{n}_{N_v^B,1}^T\right]^T, ... \left[\mathbf{n}_{1,N_r^B}^T, \mathbf{n}_{2,N_r^B}^T, ...\mathbf{n}_{N_v^B,N_r^B}^T\right]^T\right]$.

Using (1), results in the formulation of the (5) as

$$\mathbf{Y} = \sqrt{Q}\mathbf{W}^H \left(\sqrt{\frac{N_r N_v}{P}} \sum_{p=1}^{P} \psi_p \mathbf{a}^{(v)}\left(\phi_p^{(v)}\right)\left(\mathbf{a}^{(r)}\left(\phi_p^{(r)}\right)\right)^H\right) \mathbf{F} + \mathbf{N}. \tag{6}$$

Using the property of Khatri-Rao product, we can get

$$\begin{aligned}\mathbf{y} &= \sqrt{Q}\left(\mathbf{F}^T \otimes \mathbf{W}^H\right) \cdot vec\left(\mathbf{H}\right) + vec\left(\mathbf{N}\right) \\ &= \sqrt{Q}\left(\mathbf{F}^T \otimes \mathbf{W}^H\right) \cdot vec\left(\mathbf{A}_R \mathbf{H}_a \mathbf{A}_V^H\right) + \mathbf{n}_\varpi \\ &= \sqrt{Q}\left(\mathbf{F}^T \otimes \mathbf{W}^H\right) \mathbf{A}_\nu \mathbf{h}_a + \mathbf{n}_\varpi = \mathbf{Z}\mathbf{h}_a + \mathbf{n}_\varpi,\end{aligned}$$

where $\mathbf{Z} = \sqrt{Q}\left(\mathbf{F}^T \otimes \mathbf{W}^H\right)\mathbf{A}_\nu$, $\mathbf{A}_\nu = \mathbf{A}_R \otimes \mathbf{A}_V$ and $\mathbf{h}_a = vec(\mathbf{H}_a)$. Therefore, real-valued channel model is

$$\tilde{\mathbf{y}}_k = \tilde{\mathbf{Z}}_k \tilde{\mathbf{h}}_{a_k} + \tilde{\mathbf{v}}_k, \tag{7}$$

where $\tilde{\mathbf{v}}_k \in \mathbf{n}_\varpi$ is additive white Gaussian observation noise which shows no dependence on the process noise.

Moreover, the channel parameters are ψ_p, $\phi_p^{(v)}$ and $\phi_p^{(r)}$. Assuming the variation of the channel gain is very slow thus ignoring the ψ_p, results in the determination of the parameters as $\varrho_t \in \{\phi_{p_t}^{(v)}, \phi_{p_t}^{(r)}\}$ for t^{th} beam transmission period. Therefore, the discrete time model is given as $\varrho_t = \mathbf{O}_t \varrho_{t-1} + \upsilon_t$,, where $\mathbf{O}$ is the auto-regressive parameter and υ_t is the complex Gaussian vector.

B. Support Vector Regression

SVR has been adopted in diverse areas of time series prediction and has also been employed for real-time prediction [21]. The SVR method mainly constructs a linear decision function in high-dimensional space to realize model prediction. In SVR, a set of learning samples $\{(X_1, Y_1), (X_2, Y_2), ... (X_n, Y_n)\}$, the sample regression function is described as $f(X) = \omega_1 X_1 + \omega_2 X_2 + ...\omega_n X_n + b = \langle \omega, X \rangle + b$ where $\omega = [\omega_1, \omega_2, ...\omega_n]^T$ is the regression coefficient, b is the bias and $\langle\rangle$ represents the dot product. The regression problem can be solved by the constrained optimization problem [22] as min $\frac{1}{2}\|\omega\|^2 + C\sum_1^n (\xi_i + \xi_i^*)$, subject to: $Y_i - \langle \omega, X \rangle - b \leq \epsilon + \xi_i, \langle \omega, X \rangle + b - Y_i \leq \epsilon + \xi_i^*, \xi_i, \xi_i^* \geq 0$, where ξ_i and ξ^* represents slack variables to make constraints feasible, C is the regularization parameter and ϵ is the tolerance threshold. The optimization problem is easier to solve in its dual formulation, which uses the Lagrange multipliers τ_i and τ_i^*, thus gives max $-\frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n (\tau_i - \tau_i^*)(\tau_j - \tau_j^*) K(X_i, X_j) - \sum_{j=1}^n (\tau_i - \tau_i^*) Y_i + \sum_{j=1}^n (\tau_i + \tau_i^*)\epsilon$, for the condition $0 < \tau_i, \tau_i^* < C$ and $\sum_{j=1}^n (\tau_i - \tau_i^*) = 0$ and $K(X_i, X_j)$ is the kernel function.

The regression function equation can be obtained as $f(X) = \sum_{i=1}^n (\tau_i - \tau_i^*) K(\tau_i, \tau_i^*) + b$, where corresponding regression coefficient and bias can be obtained as $\langle \omega, X \rangle = \sum_{i=1}^n (\tau_i - \tau_i^*) K(\tau_i, \tau_i^*)$, $b = -\frac{1}{2}\sum_{i=1}^n (\tau_i - \tau_i^*)(K(X_i, X_r) + K(X_i, X_s))$, where X_r and X_s denote identified support vectors. In this study, the commonly used radial basis function (RBF) is selected as kernel function for SVR model and it is defined as $K(X, X_i) = exp\left(-\frac{\|X - X_i\|^2}{2\sigma^2}\right)$, where σ is the adjustable parameter to determine RBF kernel width. SVR is employed to learn motion (position) based set of training data and the trained SVR model is used to predict the future position. Therefore, in the l^{th}step look-ahead prediction, the training input vector X_i consists of m position measurements $X(i)$ at time i of the vehicle

use is given as $X_i = [X(i), X(i-1), ...X(i-m+1)]^T$. The corresponding output Y_i is $X(i+l)$ and for l^{th}step look-ahead prediction $\hat{X}_{k+l|k}$ for the input vector $X_k = [X(k), X(k-1), ...X(k-m+1)]^T$ as $\hat{X}_{k+l|k} = f(X) = \sum_{i=1}^{n} (\tau_i - \tau_i^*) K(\tau_i, \tau_i^*) + b$.

C. Unscented Kalman Filter

Owing to the non-linearity associated with the discrete time model, UKF is used for its solution, which calculates the posterior distribution of a modified random variable by extending the prior mean and covariance through the non-linear functions by way of sigma points [23]. With the initial estimate of $\hat{X}_0$ and its covariance as $\mathbf{R}_0$, the previous estimate and covariance are given as $\hat{X}_{k-1}$ and $\mathbf{R}_{k-1}$, respectively. Then the predictive state estimate for the covariance at the kth block is $\hat{X}_{k|k-1}$ and $\mathbf{R}_{k|k-1}$, respectively, given as

$$\hat{X}_{k|k-1} = \mathbf{A}_k \hat{X}_{k-1}, \tag{8}$$

with a priori covariance

$$\mathbf{R}_{k|k-1} = \mathbb{E}\left[\hat{X}_{k|k-1}\hat{X}_{k|k-1}^T\right] = \mathbf{A}_k \mathbf{R}_{k-1} \mathbf{A}_k^T + \mathbf{Q}_k. \tag{9}$$

A symmetric set of sigma point vectors $\{s_j\}_{j=1}^{2n}$ for $n = 6P$, are propagated through the transfer function $\mathscr{F}$ gives $X_j = \mathscr{F}\left(\{s_j\}_{j=1}^{2n}\right)$. Herein $\mathscr{F}$ is the representation of the channel, therefore $X_j = \tilde{\mathbf{h}}\left(\{s_j\}_{j=1}^{2n}\right)$. Therefore, the mean and covariance for $\mathbf{h}$ are approximated using a weighted sample mean and covariance of the posterior sigma points as

$$\hat{\mathbf{h}}_k = \sum_{j=0}^{2n} W_j^{(m)} \tilde{\mathbf{h}}_j, \tag{10}$$

and

$$\hat{\chi}_k = \sum_{j=0}^{2n} W_j^{(c)} \left(X_j - \hat{\mathbf{h}}\right)\left(X_j - \hat{\mathbf{h}}\right)^T, \tag{11}$$

respectively. The weights are given by $W_0^{(m)} = \frac{\lambda}{n+\lambda}$, $W_0^{(c)} = \frac{\lambda}{n+\lambda} + \left(1 - \alpha^2 + \beta\right)$ and $W_j^{(m)} = W_j^{(c)} = \frac{1}{2(n+\lambda)}, k = 1, 2, ...2n$, where $\lambda = \alpha^2(n+\gamma) - n$, for α determines the spread of the sigma points around the mean value, β is used to incorporate prior knowledge of the distribution, γ is the secondary scaling parameter [24].

Given prediction estimates $\hat{X}_{k|k-1}$ and $\mathbf{R}_{k|k-1}$, a new set of $N = 2P + 1$ sigma points $s_0, s_1, ...s_n$ with corresponding first-order weights $W_0^m, W_1^m, ...W_n^m$ and second-order weights $W_0^c, W_1^c, ...W_n^c$ is calculated [25]. These sigma points are transformed through a measurement function $\mathscr{G}$ giving $\mathbf{y}_j = \mathscr{G}(s_j), j = 0, 1, ...n$. Therefore, the empirical mean and covariance of the transformed point is given as

$$\bar{\mathbf{y}}_k = \tilde{\mathbf{y}}_k - \tilde{\mathbf{Z}}_k \tilde{\mathbf{h}}_k, \tag{12}$$

and

$$\hat{\mathbf{S}}_k = \hat{\chi}_k \tilde{\mathbf{Z}}_k \hat{\chi}_k^T + \mathbf{I}_k, \tag{13}$$

respectively, where $\mathbf{I}_k$ is the covariance matrix of the observation noise $\tilde{\mathbf{v}}_k$. Moreover, the cross covariance matrix is given as

$$\mathbf{C}_{\text{xh}} = \sum_{i=0}^{2n} W_i^{(c)} \left(X_i - \hat{X}_{k|k-1}\right)\left(\tilde{\mathbf{h}}_i - \hat{\mathbf{h}}_k\right)^T, \tag{14}$$

Therefore, the updated mean and covariance estimates are

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + \iota_g \left(\tilde{\mathbf{h}}_i - \hat{\mathbf{h}}_k\right), \tag{15}$$

and

$$\mathbf{R}_{k|k} = \mathbf{R}_{k|k-1} - \iota_g \hat{\mathbf{S}}_k \iota_g^T, \tag{16}$$

respectively, where ι_g is the Kalman gain given as $\mathbf{C}_{\text{xh}} \hat{\mathbf{S}}_k^{-1}$.

D. Federated Learning

Considering a wireless multi-user system, which consists of one server (R) and C number of local clients (vehicle users). Therefore, the total data size (T_d) for each participating client, c having a data size of d_c is given as $T_d = \sum_{c=1}^{C} d_c$. The collection of the data samples d_c is defined by the set of input-output pairs as $\{X_j, Y_j\}_{j=1}^{d_c}$, where the sample input vector is characterized by d thus $X_j \in \mathbb{R}^d$ and the output is characterized by $Y_j \in \mathbb{R}$ for sample X_j.

Therefore, for the input of X_j, a global model parameter G_m is to found out which characterizes the output Y_j, when the sampled data pair is $\{X_j, Y_j\}$. This characterization of the output with a loss function $\mathscr{L}(f_j(G_m))$ can be represented in the form: **A).** For linear regression: $\mathscr{L}(f_j(G_m)) = \frac{1}{2}\left(X_j^T(G_m) - Y_j\right)^2, Y_j \in \mathbb{R}$ and **B).** For support vector machine: $\mathscr{L}(f_j(G_m)) = \left\{0, 1 - Y_j X_j^T(G_m)\right\}, Y_j \in \{-1, 1\}$, where $X_j^T(G_m)$ is linear mapping function.

Therefore, for G_m the global optimization problem can be written as $\underset{G_m \in \mathbb{R}^d}{min} \mathscr{P}(G_m) = \sum_{c=1}^{C} \frac{d_c}{T_d} \mathscr{L}(\mathscr{P}_c(G_m))$, where for the user u the local loss function is given as $\mathscr{L}(\mathscr{P}_c(G_m)) = \frac{1}{d_c} \sum_{j=d_c} \mathscr{L}(f_j(G_m))$, and the loss function $\mathscr{L}(f_j(G_m))$ captures the error of the model parameter G_m on the input-output pair $\{X_j, Y_j\}$. Along with it, its solution is captured by iterative FL-based approach and each iterative round is indexed by I, and it consists of three steps:

1) **Scheduling of participants:** The main server, R selects the initial model parameters G_{m_p} and schedules R_d for its training and then broadcasts these details to each client participant $C_{p_i}, i \in 1, 2..N_c$, where N_c are a number of clients for an iterative round I. C_{p_i} receives the current global model parameter, $G_{m_{p_{I-1}}}^{S_{d_{I-1}}}$ from the S on the past scheduling decisions given as $R_{d_{I-1}} = [R_{d_1}, R_{d_2}, ...R_{d_{I-1}}]$.
2) **Schedule sharing:** Each C_{p_i} has its own local data set, thus using the gradient descent algorithm updates its local model parameter, $\tilde{G}_{m_p} \in G_{m_p}$ for a learning rate of μ, as $\tilde{G}_{m_p}(I+1) = \tilde{G}_{m_p}(I) - \mu \tilde{\nabla} f_j\left(G_{m_p}(I)\right)$, where for a mini-batch $d_{C_p} \subseteq D_{C_p}$ of $C_p \in c$ data samples, the estimate of the gradient $\nabla f_j\left(G_{m_p}(I)\right)$ is given as $\tilde{\nabla} f_j\left(G_{m_p}(I)\right) = \frac{1}{|d_{C_p}|} \sum_{j \in d_{C_p}} \nabla \mathscr{L}(f_j\left(G_{m_p}(I)\right))$.

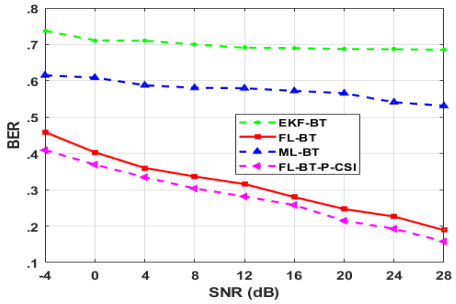

Fig. 2: Effect of SNR on BER

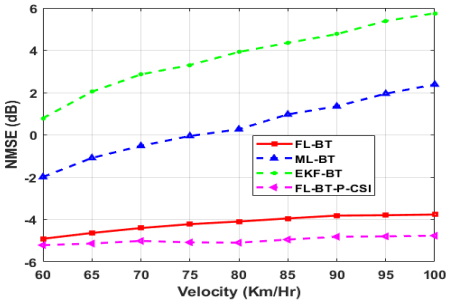

Fig. 3: Effect of Velocity on NMSE

The exchange of the local model parameter is scheduled and repeated α times, represented as $f_j(G_{m_p}(\alpha))$ is uploaded and shared to R.

3) **Update global model:** Weighted average combination of the local parameters received is carried out by the R, which aggregates them to obtain a new global model parameter, which is represented as $G_{m_p}(I)^{R_{d_I}} = \frac{\sum_{c\in R_{d_I}} d_c G_{m_p}(I)^{R_{d_I}}}{\sum_{c\in R_{d_I}} d_c}$.

IV. Performance Evaluation

A frequency band of 28 GHz with ULAs is considered, with communication between a single RSU and a single VU. Along with it, BS and VU are equipped with 32 transmitting and receiving antennas, respectively. In order to predict the channel vector of vehicular users, the speed range of vehicle users is considered from 60 km/h to 100 km/h. As the performance measure normalized mean square error (NMSE) of $\mathbf{H}$ which is given as $\frac{\mathbb{E}\left(\left\|\hat{\mathbf{H}}-\mathbf{H}\right\|_F^2\right)}{\mathbb{E}\left(\|\mathbf{H}\|_F^2\right)}$, where $\mathbf{H}$ is the true channel matrix, $\hat{\mathbf{H}}$ is the estimated channel matrix and $\|\cdot\|_F$ is the Frobenius norm. The accusation of the following mobility parameters: velocity, acceleration, angular velocity, and angular acceleration is based on the approach presented by [26], details of which are not discussed in this paper.

Fig. 2 shows the bit error rate (BER) performance against the signal-to-noise ratio (SNR) at an average speed of vehicle user 70 km/h, in which the proposed approach, FL-BT, not only outperforms the existing methods, but also consistently pursues FL-BT-P-CSI. This further compounds the perseverance of the proposed approach. The BER not only starts from the lowest value for FL-BT but also decreases almost exponentially to the lowest values as compared to contemporary approaches, which shows a decrease that is almost negligible. This demonstrates that a more accurate model is provided by the SVR-based channel model, which is further supported by the UKF, thereby compensating for the model's non-linearities. As a result, better performance is obtained under conditions of increased mobility and inaccurate CSI.

Fig. 3 illustrates the variation in NMSE performance with the vehicle user's velocity. All algorithms' performance suffers when the channel gets more dynamic because vehicle user velocity rises. But this degrading behavior for FL-BT is marginal, whereas with the existing approach, significant differences can be observed, and moreover, it follows FL-BT-P-CSI very closely, thereby proving its fortitude even in the presence of imperfect CSI. The linear channel model used in the EKF approach does not adequately describe the behavior of time-varying channels; hence, EKF-BT's performance is poor compared to other methods. In contrast, the suggested solution for accurate beam tracking successfully predicts complex channel behavior.

The behavior of the NMSE against the SNR is compared and presented in Fig. 4 for all three approaches. A significant gain in performance is observed for the proposed approach as compared to existing methods, and an impending FL-BT-P-CSI is anticipated. This endorses the efficient behavior of FL-BT. The proposed method performs better than EKF because its prediction results at the prediction update stage are more accurate. This is further enhanced by the proposed method's federated structure, as a result of which it can follow the

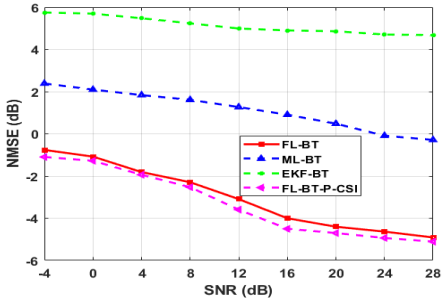

Fig. 4: Effect of SNR on NMSE

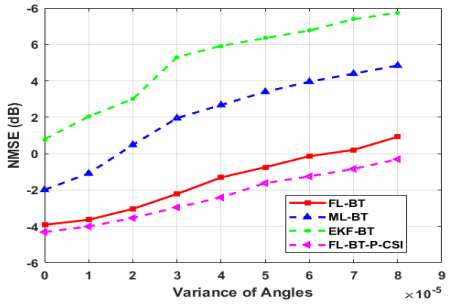

Fig. 5: Effect of Variance of Angles on NMSE

channel's quick variations well.

Fig. 5 shows and compares the performance of the proposed approach with contemporary methods, in terms of NMSE against the variance of the angles (for AoAs). As the variance starts to increase, ML-BT and EKF-BT lose track, which leads to an increase in the estimation error. The proposed approach compensates for the degraded channel estimation, thereby giving exceptional performance by maintaining a course that is adjacent to FL-BT-P-CSI.

The average estimation error for NMSE is compared between the proposed approach, EKF-BT and ML-BT, Fig. 6, for variation in the number of paths. For the FL-BT, the maximum value is almost 50% lower than the other two approaches and the mean value of the average NMSE error for the proposed approach is 0.0037, for ML-BT it is 0.0.0136 and for EKF-BT is 0.0.0406, this clearly substantiate that SVR and UKF with the augment of FL is able to lower the NMSE.

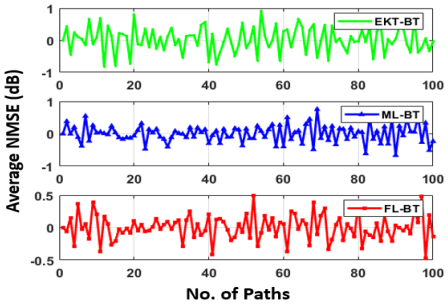

Fig. 6: Effect of No. of Paths on NMSE

V. Conclusion

This paper proposed FL-BT, which stands for federated learning-based combined radar-communication mmWave beamtracking for V2X communication under poor CSI. Because the vehicle user is moving, the proposed beamtracking approach was created to track rapidly varying AoD and AoA. To calculate channel variations and forecast future channel state estimates, the SVR and UKF-based federated module for the vehicle user and RSU are used. Our simulation findings demonstrated that the suggested method outperformed the current beamtracking methods in the presence of suboptimal CSI and achieved a sizable performance boost.

References

[1] Y. Liu and C. Tang, "Concurrent multi-beam transmissions for reliable communication in millimeter-wave networks," *Computer Communications*, vol. 195, pp. 281–291, 2022.

[2] J. Zhang, Y. Huang, J. Wang, X. You, and C. Masouros, "Intelligent interactive beam training for millimeter wave communications," *IEEE Transactions on Wireless Communications*, vol. 20, no. 3, pp. 2034–2048, 2020.

[3] Z. Wei, F. Zhang, S. Chang, Y. Liu, H. Wu, and Z. Feng, "MmWave Radar and Vision Fusion for Object Detection in Autonomous Driving: A Review," *Sensors*, vol. 22, no. 7, p. 2542, 2022.

[4] J. Choi, V. Va, N. Gonzalez-Prelcic, R. Daniels, C. R. Bhat, and R. W. Heath, "Millimeter-wave vehicular communication to support massive automotive sensing," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 160–167, 2016.

[5] P. Kumari, N. J. Myers, and R. W. Heath, "Adaptive and fast combined waveform-beamforming design for mmWave automotive joint communication-radar," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 4, pp. 996–1012, 2021.

[6] Y. Guo, Z. Wang, M. Li, and Q. Liu, "Machine learning based mmwave channel tracking in vehicular scenario," in *2019 IEEE International conference on communications workshops (ICC Workshops)*. IEEE, 2019, pp. 1–6.

[7] A. P. Hermawan, R. R. Ginanjar, D.-S. Kim, and J.-M. Lee, "CNN-based automatic modulation classification for beyond 5G communications," *IEEE Communications Letters*, vol. 24, no. 5, pp. 1038–1041, 2020.

[8] T. Huynh-The, Q.-V. Pham, T.-V. Nguyen, T. T. Nguyen, R. Ruby, M. Zeng, and D.-S. Kim, "Automatic modulation classification: A deep architecture survey," *IEEE Access*, vol. 9, pp. 142 950–142 971, 2021.

[9] C. Zhang, D. Guo, and P. Fan, "Tracking angles of departure and arrival in a mobile millimeter wave channel," in *2016 IEEE international conference on communications (ICC)*. IEEE, 2016, pp. 1–6.

[10] A. Barrau, "Non-linear state error based extended Kalman filters with applications to navigation," Ph.D. dissertation, Mines Paristech, 2015.

[11] N. Veerakumar, D. Četenović, K. Kongurai, M. Popov, A. Jongepier, and V. Terzija, "PMU-based Real-time Distribution System State Estimation Considering Anomaly Detection, Discrimination and Identification," *International Journal of Electrical Power & Energy Systems*, vol. 148, p. 108916, 2023.

[12] S. J. Julier and J. K. Uhlmann, "New extension of the Kalman filter to nonlinear systems," in *Signal processing, sensor fusion, and target recognition VI*, vol. 3068. Spie, 1997, pp. 182–193.

[13] S. S. Haykin and S. S. Haykin, *Kalman filtering and neural networks*. Wiley Online Library, 2001, vol. 284.

[14] F. Gustafsson, "Particle filter theory and practice with positioning applications," *IEEE Aerospace and Electronic Systems Magazine*, vol. 25, no. 7, pp. 53–82, 2010.

[15] P. Kumari and D. Toshniwal, "Extreme gradient boosting and deep neural network based ensemble learning approach to forecast hourly solar irradiance," *Journal of Cleaner Production*, vol. 279, p. 123285, 2021.

[16] Q. M. Ilyas, K. Iqbal, S. Ijaz, A. Mehmood, and S. Bhatia, "A Hybrid Model to Predict Stock Closing Price Using Novel Features and a Fully Modified Hodrick–Prescott Filter," *Electronics*, vol. 11, no. 21, p. 3588, 2022.

[17] J. D. Borrero, J. Mariscal, and A. Vargas-Sánchez, "A New Predictive Algorithm for Time Series Forecasting Based on Machine Learning Techniques: Evidence for Decision Making in Agriculture and Tourism Sectors," *Stats*, vol. 5, no. 4, pp. 1145–1158, 2022.

[18] P. Fiszeder and W. Orzeszko, "Covariance matrix forecasting using support vector regression," *Applied intelligence*, vol. 51, no. 10, pp. 7029–7042, 2021.

[19] Y. You and L. Zhang, "Exploiting angular spread in channel estimation of millimeter wave MIMO system," *IET Communications*, vol. 16, no. 3, pp. 195–205, 2022.

[20] A. Ghasemi and S. R. Zekavat, "On Eigenvalue Distribution of Imperfect CSI in mmWave Communications," in *2022 IEEE USNC-URSI Radio Science Meeting (Joint with AP-S Symposium)*. IEEE, 2022, pp. 56–57.

[21] Y. Bao, T. Xiong, and Z. Hu, "Multi-step-ahead time series prediction using multiple-output support vector regression," *Neurocomputing*, vol. 129, pp. 482–493, 2014.

[22] V. Vapnik, *The nature of statistical learning theory*. Springer science & business media, 1999.

[23] R. Van Der Merwe, *Sigma-point Kalman filters for probabilistic inference in dynamic state-space models*. Oregon Health & Science University, 2004.

[24] K. György, A. Kelemen, and L. Dávid, "Unscented Kalman filters and Particle Filter methods for nonlinear state estimation," *Procedia Technology*, vol. 12, pp. 65–74, 2014.

[25] S. Sarkka, "On unscented Kalman filtering for state estimation of continuous-time nonlinear systems," *IEEE Transactions on automatic control*, vol. 52, no. 9, pp. 1631–1641, 2007.

[26] S. H. Lim, S. Kim, B. Shim, and J. W. Choi, "Deep Learning-based Beam Tracking for Millimeter-wave Communications under Mobility," *IEEE Transactions on Communications*, vol. 69, no. 11, pp. 7458–7469, 2021.