MESSAGE

IEICE Transactions on Communications: Editor's Message

The month of June marks the beginning of IEICE's fiscal year when the Editorial Board of IEICE Transactions will start with new succeeding members as some board members have completed their terms of office. With that, it is an honor and privilege to announce that I will now serve as the Editor-in-Chief (EiC), for two years, of the IEICE Transactions on Communications.

I am compelled to reflect on the profound societal shifts precipitated by the COVID-19 pandemic. This unprecedented global crisis has acted as a catalyst for the rapid acceleration of digital transformation across various sectors, fundamentally altering the way we live, work, and interact. As the world grapples with the complexities of navigating a post-pandemic landscape, it becomes increasingly evident that information and communication technologies have emerged as indispensable pillars supporting resilience, innovation, and progress.

Having a 105-year history since its establishment in 1917, the IEICE contributes to the advancement of information and communication technology, and the IEICE Transactions on Communications has been serving as one of the major platforms for the dissemination of technologies developed by national and international researchers and engineers since it was first issued in 1991. However, since other journals are also increasingly taking the role of new and innovative technology dissemination, the IEICE Transaction needs to increase its competitiveness. To achieve this, we will continuously work to improve our services, such as shortening the review period, introducing advanced publications, position paper categories, and accepting papers submitted by non-IEICE members. In addition, to maintain reasonable review and decision-making, which is most important for authors, review criteria are being checked by sharing case studies with all members of the editorial committee at every quarter meeting. As a result of these steady efforts, the number of submissions last year increased by around 20% compared with that in the previous year. Accessibility to published articles is another key issue for academic journals, in which the former EiC, Prof. Cho has devoted his efforts to tackle this issue and realized the migration of the IEICE Transactions on Communications to IEEE Xplore. From now on, it is a privilege to take up the baton, and would like to push forward with the launch of other novel plans. I will do my best to continuously increase the value of the transaction with my fellow Editors, Prof. Takeshi Amishima, Prof. Tohru Kondo, Prof. Kazuya Tsukamoto, and all the committee members.

Finally, I would like to express my deepest gratitude to Prof. Keizo Cho, the former EiC, Prof. Katsuyoshi Iida, the former Editor, and the former associate editors, for their great effort to enhance the competitiveness of this transaction. Last, but not least, I would like to give my special thanks to the authors, readers, and reviewers for their significant contributions in submitting, reviewing, and citing their outstanding works, which are of the utmost importance to the value of this transaction.

Go Hasegawa, Editor-in-Chief

Go Hasegawa *(Member)* a member of the IEEE, received the ME and DE degrees in information and computer sciences from Osaka University, Japan, in 1997 and 2000, respectively. From July 1997 to June 2000, he was a research assistant at the Graduate School of Economics, Osaka University. From 2000 to 2018, he was an associate professor with the Cybermedia Center at Osaka University. He is now a professor at the Research Institute of Electrical Communication, Tohoku University. His research is in information network architecture, wireless networking architecture, edge-cloud computing, Internet congestion control, and so on. He has served as Associate Editor and Editor of the IEICE Transactions on Communications for 10 years and has received the Best Paper Award from the IEICE in 2010.

IEICE TRANS. COMMUN., VOL.E107–B, NO.6 JUNE 2024

PAPER

Physical Layer Security Enhancement for mmWave System with Multiple RISs and Imperfect CSI

Qingqing TU[†*a)], ***Member***, **Zheng DONG**[††], **Xianbing ZOU**[†], ***and*** **Ning WEI**[†], ***Nonmembers***

SUMMARY Despite the appealing advantages of reconfigurable intelligent surfaces (RIS) aided mmWave communications, there remain practical issues that need to be addressed before the large-scale deployment of RISs in future wireless networks. In this study, we jointly consider the non-neglectable practical issues in a multi-RIS-aided mmWave system, which can significantly affect the secrecy performance, including the high computational complexity, imperfect channel state information (CSI), and finite resolution of phase shifters. To solve this non-convex challenging stochastic optimization problem, we propose a robust and low-complexity algorithm to maximize the achievable secrete rate. Specially, by combining the benefits of fractional programming and the stochastic successive convex approximation techniques, we transform the joint optimization problem into some convex ones and solve them sub-optimally. The theoretical analysis and simulation results demonstrate that the proposed algorithms could mitigate the joint negative effects of practical issues and yielded a trade-off between secure performance and complexity/overhead outperforming non-robust benchmarks, which increases the robustness and flexibility of multiple RIS deployments in future wireless networks.

key words: *reconfigurable intelligent surfaces (RIS), physical layer security, millimeter wave (mmWave), imperfect CSI, robust beamforming*

1. Introduction

The millimeter wave (mmWave) communications can significantly boost network capacity to make ubiquitous and on-demand interconnection possible in future wireless networks [1]–[3]. However, owing to the propagation loss of the mmWave band and the broadcast nature of wireless channels, mmWave communication suffers from the risk of information leakage in the presence of eavesdroppers, especially when the legitimate and wiretap channels are highly correlated, which makes security an indispensable pursuit for system development. In this context, physical layer security (PLS) has drawn significant research interest in recent years [4], [5]. Compared with traditional cryptography, the PLS techniques enable secrecy communication independently of the higher layers by taking advantage of the inherent randomness of noise and communication channels, such as cooperative relaying schemes, artificial noise-aided beamforming, and cooperative jamming [6]–[9]. However, enhancing the PLS by deploying a large number of relays or other active nodes will inevitably increase the cost and complexity of the communication systems. To solve this problem, reconfigurable intelligent surfaces (RIS) have recently been introduced to provide a new paradigm for the PLS enhancement [10], [11]. Unlike active components that formulate active beamforming to transmit signal, such as base stations, the RIS achieves passive beamforming through its passive planar surface. This surface consists of a large number of passive reflecting elements, each capable of inducing a controllable amplitude [12] and/or phase [13] changes to the incident signal. Thus, it incurs lower energy costs compared to active antenna arrays. With this property, the RIS can serve as an auxiliary device flexibly deployed in wireless networks to boost or suppress the received signals and also improve the network coverage by creating virtual LOS links, especially in the higher frequency bands like mmWave [14] to deal with the severe path loss and the signal blockage problem. These unique merits stimulate the increasing interest in RIS-aided PLS enhancement research. As illustrated in [15]–[17], the secure performance is investigated in the RIS-aided wireless communication systems with both legitimate receivers and eavesdroppers existing. By developing various non-convex optimization techniques, such as the successive convex approximation (SCA)-based method, the semidefinite relaxation (SDR)-based method, and the manifold optimization (MO)-based method, it revealed that the secure performance of the system with RIS outperforms the case without RIS, which proves the advantage of deployment of the RIS in the secure transmission.

However, the low-rank channel of a single RIS to AP is hard to support secure transmissions in the existence of eavesdroppers, especially when the wiretap channel is stronger than the legitimate user channel. To conquer this issue, multiple RIS deployment is investigated to further enhance the secure performance of ubiquitous wireless networks [18], [19]. Furthermore, the benefits brought by the multiple RIS deployment also perfectly match with the secure mmWave communications. To circumvent the inherent disadvantages of mmWave, typical multiple-input multiple-output (MIMO) technologies with massive array antennas are deployed to cope with the high path-loss problem [20]–[22]. However, continuous increment in antenna number is impractical due to the high hardware complexity and power consumption [23] and the blockage-prone issue has still not been well addressed. A more feasible solution is to in-

Manuscript received October 21, 2023.
Manuscript publicized January 30, 2024.
[†]The authors are with the National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 611731, China.
[††]The author is with the School of Information Science and Engineering, Shandong University, Qingdao, 266237, Shandong Province, China.
[*]Presently, with the National Computer Network Emergency Response Technical Team/Coordination Center of China.
a) E-mail: tuqingqing@cert.org.cn
DOI: 10.23919/transcom.2023EBP3169

troduce distributed RISs to aid secure communication in the mmWave system, which provides a more cost-efficient way to compensate for the limited secure propagation distance problem by forming virtual line-of-sight (LoS) paths as proved in [24], [25]. By jointly optimizing active and passive beamformers, the author in [26] maximized the secrecy rate and proved the significant secure performance gains offered by the multiple RISs deployment when the eavesdropper exists. However, the multi-RIS-aided PLS enhancement performance for the mmWave system is still restricted by practical issues.

The first challenging issue lies in the acquisition of accurate channel state information (CSI) of the RIS due to its passive operation and a large number of reflecting elements. In general, the performance of the joint optimization for passive beamforming and active beamforming is highly dependent on the quality of the acquired CSI. Despite the various methods for the RIS channel estimation [27]–[29], the estimation error is still inevitable because of the channel background noise, time-varying characteristics, and more importantly, the fundamental limitation of no active transmitting elements on the RIS [30]. Although the previous works on RIS aided secure transmission treating the estimated channels as perfect ones achieved encouraging results, it will inevitably lead to system performance loss and weaken the generality and practicality of the systems and algorithms. Hence, it is crucial to consider the CSI uncertainties for the RIS-aided PLS enhancement scheme. In the recent work to deal with the imperfect CSI problem, one way is to introduce artificial noise (AN) to deliberately destroy the wiretap channel and reduce the dependency on CSI by optimizing the power fraction of AN and the RIS phase shifts [31]–[33]. However, the extra power is located to transmit the AN signal, which is not cost-efficient. Another way is to use the imperfect CSI to design the joint beamforming in the RIS-aided system. To model the uncertain CSI error, the bounded error model is commonly used in existing works [34]–[36], which proposed a worst-case robust beamforming design and obtained a suboptimal solution by setting the channel quantization error within a bounded region. However, this model may not adapt to apply in the realistic channel, since the estimation error is Gaussian due to the linear channel estimator under Gaussian noise. This problem can be handled in the Gaussian CSI error model (or statistical CSI error model), which could introduce stochastic optimization in the secure design for RIS-aided networks. Specifically, the authors in [30] showed that better performance could be obtained in a robust beamforming design under this type of CSI error model in terms of convergence speed and complexity. To make the secure transmission design more practical, the authors of [37]–[40] investigated different performance metrics for RIS-aided systems under imperfect CSI of eavesdroppers, such as the average secrecy rate, energy efficiency, and secrecy outage probability, which proposed different algorithms to improve the security performance. However, we note that all these secure works only considered the scenarios with the imperfect CSI of the eavesdroppers, while no relevant research on analyzing the secure performance of multi-RIS-aided systems considering imperfect CSI of both intended users and eavesdroppers, which is more practical. Moreover, compared with the previous work, this optimization problem is more difficult to solve, and the converged solution under alternating optimization iteration is uncertain due to the random system states introduced by the imperfect CSI of both the intended user and the eavesdropper. Hence, due to the lack of adequate material in the literature, the secure performance of multiple RISs in mmWave systems under such imperfect CSI cases is yet to be explored and utilized up to their full potential.

The second non-neglectable practical issue is the high computational complexity in multi-RIS-aided systems due to the large-scale reflection coefficients optimization [41]. The number of reflection elements on RIS is typically much larger than that of the antennas on the access point (AP), and the reflection coefficients optimization on the RISs is computationally prohibitive, especially when multiple RISs are deployed. In addition, the complex methods applied to address the non-convex joint optimization problem result in increased complexity, such as the SDR-based method [13], SCA-based method [42], and MO-based method [26], [43], which will further restrict the realization of the full potential of RIS. Thus, reducing the computational complexity of joint beamforming design is essential for the practical deployment of multiple RISs.

The third practical issue is the adoption of phase shifters with a finite resolution on the RIS to strike a balance between hardware cost and system performance. For simplicity, many aforementioned RIS-aided security schemes implicitly assume that infinite-resolution phase shifters are available on the RIS, which are arguably prohibitive to implement [44]. On the other hand, applying finite-resolution phase shifters will inevitably incur a notable secrecy performance loss [41]. Previous studies have focused on addressing issues such as transmitted power minimization, sum-rate maximization, and coverage improvement using the finite-resolution phase shifters in the specific terminal settings [45]–[48], which encompass parameters like RIS size, power constraint, rate constraint. The work in [48] further investigated the required number of phase shifts under a rate constraint. Nevertheless, the RIS-aided PLS enhancement scheme under the finite-resolution phase shifters case still needs further investigation.

In summary, with the increased demands for the RIS deployment scale in the future, it is worth extending the secure transmission designs to more practical setups in the mmWave systems. However, jointly considering these non-neglectable practical issues leads to a more challenging stochastic optimization due to more coupled variables, complex objective functions, non-convex constraints, and random system states. To our best knowledge, this is still an open problem and requires new solution approaches.

In this paper, we consider an mmWave system with multiple RISs deployed against eavesdroppers and investi-

gate the secrecy performance gain jointly affected by these practical issues. To address this challenging non-convex stochastic optimization problem, we propose a robust and low-complexity method to provide trade-offs concerning secrecy performance, complexity, and cost. The main contributions of this paper are summarized as follows.

- We formulate a robust secure beamforming problem to maximize the worst-case achievable SR under the joint effects of non-neglectable practical issues. Different from the existing RIS-aided work under the imperfect CSI of eavesdroppers [37]–[40], our design simultaneously considers the imperfect CSI of both intended users and eavesdroppers and also expanded to the low-resolution shifters case.
- To solve this challenging non-convex stochastic optimization problem, we expressed the imperfect CSI model with the random system states and propose robust and low-complexity methods to transform the objective problem into simplified and convex subproblems by combining the benefits of the fractional programming and the stochastic successive convex approximation techniques. We mathematically prove that the convergence conditions are satisfied in each subproblem optimization, and the resulting achievable SR under alternating optimization iteration is guaranteed to converge.
- The proposed method exhibits a robustness improvement over the non-robust benchmarks and achieves a close secure performance gain to that of perfect CSI performance when the estimation error is bounded, while at a lower computational complexity via theoretical analysis and simulations. Furthermore, the RIS of 3-bit phase adjustable elements achieves the equivalent value as the continuous phase shift. These proved that the proposed method can reduce the demand for perfect CSI and infinite hardware resolution, which increases the flexibility and robustness of multiple RIS deployment in various practical scenarios.

Notations: Variables, vectors, and matrices are respectively written as lower-case letters, bold lower-case letters and bold upper-case letters. $\mathbf{A}_{m,n}$ is the the m-th row and n-th column element of the matrix $\mathbf{A}$. $\mathrm{tr}(\cdot)$, $(\cdot)^T$, and $(\cdot)^H$ stand for the trace, the transpose, and the conjugate transpose, respectively, while $(\cdot)^*$ donates the conjugate. $\|\cdot\|$ and $\|\cdot\|_F$ is the Euclidean norm and the Frobenius norm of a complex vector and matrix, respectively. $\mathrm{diag}(\cdot)$ represents the diagonal matrix whose diagonals are the elements of the input vector. j denotes the imaginary unit. $[\cdot]_{(1:m)}$ returns the vector that contains the first m elements. Finally, $\log(x)$ represents base 2 logarithm of x.

2. System Model and Problem Formulation

This section describes a multi-RIS-aided downlink mmWave system with both the intended user and eavesdropper. In addition, the channel model, which considers the CSI uncertainties, is also discussed.

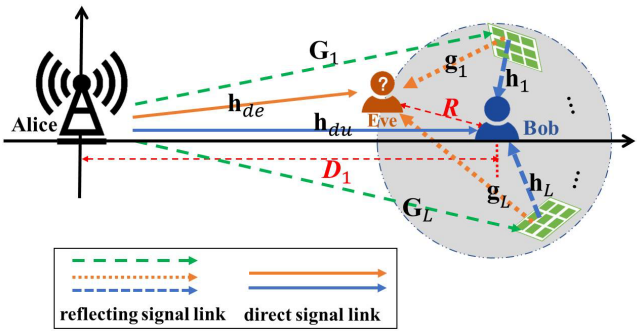


Fig. 1 The multi-RIS-aided secure mmWave system.

2.1 System Model

We consider the downlink of a TDD-based multi-RIS-aided mmWave system where the eavesdropper (Eve) attempts to eavesdrop on the confidential message sent from an access point (Alice) to the receiver (Bob), which is under a worst-case assumption that near-end Eve tries to intercept the information of far-end Bob, which could lead to the zero secrecy rate (SR) problem [41]. Specially, we assume that both Eve and Bob are regular users of the system and hence know the communication protocols. In addition, a set of L RISs are distributively deployed [49] to enhance the SR as well as the network coverage by reflecting the received signals, as depicted in Fig. 1.

We assume that Alice has N antennas, and each RIS is equipped with $M = M_1 \times M_2$ reflecting elements arranged in a uniform planer array (UPA) with M_1 elements horizontally and M_2 elements vertically. Following [41], the phase shifter matrix of the l-th RIS is $\mathbf{\Theta}_l = \mathrm{diag}(\boldsymbol{\theta}_l) \in \mathbb{C}^{M\times M}$, $\boldsymbol{\theta}_l = \sqrt{\eta}[\theta_{1,l}, \cdots, \theta_{M,l}]^T$, in which η is the amplitude reflection coefficient, and $\theta_{m,l} = e^{j\phi_{m,l}}, m = 1, \cdots, M, l \in \mathcal{L}$. We can adjust the phase shifts $\phi_{m,l} \in [0, 2\pi)$ of each reflecting element to create a virtual line-of-sight (LoS) link.

The channel matrix between Alice and the l-th RIS is represented by $\mathbf{G}_l \in \mathbb{C}^{M\times N}$. The channel vector between the l-th RIS and Bob is denoted by $\mathbf{h}_l \in \mathbb{C}^{M\times 1}$, while $\mathbf{g}_l \in \mathbb{C}^{M\times 1}$ is the channel vector between the l-th RIS and Eve. The LoS channel vectors, the Alice-Bob link and the Alice-Eve link, are represented by $\mathbf{h}_{du} \in \mathbb{C}^{N\times 1}$ and $\mathbf{h}_{de} \in \mathbb{C}^{N\times 1}$, respectively. Now, the received signal at Bob over the cascaded channel via the RIS can be modeled by

$$y_u = \left(\sum_{l=1}^{L} \mathbf{h}_l^H \mathbf{\Theta}_l \mathbf{G}_l + \mathbf{h}_{du}^H\right) \mathbf{w} s + n, \tag{1}$$

where s is the confidential transmitted signal such that $\mathbb{E}[|s|^2] = 1$, and $\mathbf{w} \in \mathbb{C}^{N\times 1}$ is the beamforming vector subject to a total transmit power constraint $\mathrm{tr}(\mathbf{w}\mathbf{w}^H) \le P$. In addition, $n \sim \mathcal{CN}(0, \sigma^2)$ is an additive zero-mean Gaussian noise at Bob. Likewise, the signal received at Eve can be given by

$$y_e = \left(\sum_{l=1}^{L} \mathbf{g}_l^H \mathbf{\Theta}_l \mathbf{G}_l + \mathbf{h}_{de}^H\right)\mathbf{w}s + n_e, \tag{2}$$

where $n_e \sim \mathcal{CN}(0, \sigma_e^2)$ is the additive complex Gaussian noise at Eve. The signal-to-interference-plus-noise ratio (SINR) of the received signal of Bob is formulated by

$$\gamma_u = \frac{\|(\sum_{l=1}^{L} \mathbf{h}_l^H \mathbf{\Theta}_l \mathbf{G}_l + \mathbf{h}_{du}^H)\mathbf{w}\|^2}{\sigma^2}. \tag{3}$$

The SINR for Eve is given in a similar fashion by

$$\gamma_e = \frac{\|(\sum_{l=1}^{L} \mathbf{g}_l^H \mathbf{\Theta}_l \mathbf{G}_l + \mathbf{h}_{de}^H)\mathbf{w}\|^2}{\sigma_e^2}. \tag{4}$$

Therefore, achievable SR between Alice and Bob can be written as

$$R_s(\mathbf{w}, \mathbf{\Theta}) = \left[\log(1+\gamma_u) - \log(1+\gamma_e)\right]^+, \tag{5}$$

where $[x]^+ \triangleq \max\{0, x\}$. It is worth pointing out that the optimal value of R_s is always non-negative by adjusting the beamforming vectors. This can be shown by contradiction. If we assume that the optimal value $R_s(\mathbf{w}^*, \mathbf{\Theta}^*)$ is negative and $\mathbf{w}^*$ is the optimal value. Then, by setting $\|\mathbf{w}\| = 0$, we can obtain $R_s = 0$, which contradicts our assumption.

2.2 Channel Model

In our system, there are two types of channels: the direct channel between Alice and Bob (or Eve) $\mathbf{h}_{d,i}, i = \{u, e\}$, and the RIS-aided cascaded channel $\mathbf{G}$, $\mathbf{g}$, and $\mathbf{h}$. We assume that the direct channel $\mathbf{h}_{d,i}, i = \{u, e\}$ with rich scatters are Rayleigh distributed, while the channel $\mathbf{G}$, $\mathbf{g}$, and $\mathbf{h}$ follow a Rician fading, since the LoS components are generally contained in the RIS-aided cascaded channel, which is similar to the model described in [40], [43], [50]. More specifically, the Alice-to-RIS, the RIS-to-Bob, and the RIS-to-Eve mmWave channels are expressed as followed

$$\mathbf{G} = \left(\sqrt{\frac{\kappa_1}{\kappa_1+1}}\bar{\mathbf{G}} + \sqrt{\frac{1}{\kappa_1+1}}\hat{\mathbf{G}}\right)L_{AR}, \tag{6a}$$

$$\mathbf{h} = \left(\sqrt{\frac{\kappa_2}{\kappa_2+1}}\bar{\mathbf{h}} + \sqrt{\frac{1}{\kappa_2+1}}\hat{\mathbf{h}}\right)L_{RU}, \tag{6b}$$

$$\mathbf{g} = \left(\sqrt{\frac{\kappa_3}{\kappa_3+1}}\bar{\mathbf{g}} + \sqrt{\frac{1}{\kappa_3+1}}\hat{\mathbf{g}}\right)L_{RE}, \tag{6c}$$

where $\bar{\mathbf{G}}$, $\bar{\mathbf{h}}$ and $\bar{\mathbf{g}}$ are the corresponding LoS components, which remains unchanged within the channel coherence time. The Rician factors in each channel are represented by $\kappa_1, \kappa_2, \kappa_3$, respectively. Also, $\hat{\mathbf{G}}, \hat{\mathbf{h}}, \hat{\mathbf{g}}$ are the channel error terms corresponding to the non-line-of-sight (NLoS) components, and each of their elements is i.i.d. CSCG random variable with zero mean and unit variance. The path-loss terms are captured by L_{AR}, L_{RU}, and L_{RE}.

Based on the mmWave channel model in [51], The LoS component of the channel between the Alice and the RIS with L_G scatters can be formulated by

$$\bar{\mathbf{G}} = \frac{1}{\sqrt{L_G}}\sum_{l_G=1}^{L_G} \mathbf{u}_t\left(\varphi_{l_G}^t\right)\mathbf{u}_r\left(\psi_{l_G}^r, \varphi_{l_G}^r\right). \tag{7}$$

The array response vector at Alice is denoted by $\mathbf{u}_t$ such that

$$\mathbf{u}_t(\varphi_{l_G}^t) = \frac{1}{\sqrt{N}}\left[1, e^{j\frac{2\pi d}{\lambda_c}\sin\left(\varphi_{l_G}^t\right)}, \cdots, e^{-j(N-1)\frac{2\pi d}{\lambda_c}\sin\left(\varphi_{l_G}^t\right)}\right], \tag{8}$$

where λ_c is the wavelength, d is the antenna spacing, and $\varphi_{l_G}^t \in [0, 2\pi)$ represents the angle of departure. For a typical UPA, array response at RIS $\mathbf{u}_r$ is expressed as

$$\begin{aligned}\mathbf{u}_r(\psi_{l_G}^r, \varphi_{l_G}^r) = &\frac{1}{\sqrt{M}}\left[1, \cdots, e^{-j\frac{2\pi d}{\lambda_c}(M_1-1)\cos\left(\varphi_{l_G}^r\right)\sin\left(\psi_{l_G}^r\right)}\right] \\ &\otimes\left[1, \cdots, e^{-j\frac{2\pi d}{\lambda_c}(M_2-1)\cos\left(\psi_{l_G}^r\right)}\right],\end{aligned} \tag{9}$$

where $\psi_{l_G}^r$ and $\varphi_{l_G}^r \in [0, 2\pi]$ represent the elevation and azimuth angle of arrival at the l-th RIS, respectively. The cascaded channel between the RISs to Bob $\mathbf{h}$ and Eve $\mathbf{g}$ is defined similarly and omitted for simplicity.

With the limited propagation range due to the strong path loss and fading effects encountered by millimeter-wave signals during transmission, the spatial correlation between the eavesdropping channel and the legitimate channel significantly influences the security performance of the system, particularly when the eavesdropper is in close proximity to the legitimate communication devices. Therefore, we introduce a correlation coefficient to describe the different spatial correlations on the receiver sides of Bob and Eve following [52], [53], which is expressed as

$$\rho_r = \frac{\mathbb{E}\{\mathbf{h}_{du}^H \mathbf{h}_{de}\}}{\sqrt{\mathbb{E}\{\mathbf{h}_{du}^H \mathbf{h}_{du}\}\mathbb{E}\{\mathbf{h}_{de}^H \mathbf{h}_{de}\}}}. \tag{10}$$

Therefore, the RIS-to-Bob and the RIS-to-Eve mmWave channels can be expressed as $\mathbf{h}_{d,i} = \sqrt{\rho_r}\mathbf{h}_{d,0} + \sqrt{1-\rho_r}\mathbf{h}_{d,i,1}$. $\mathbf{h}_{d,0}$ is the correlated component referring to the deterministic part of the channels. For $\rho_r \in [0, 1]$, $\mathbf{h}_{d,i,1}$ donates the uncorrelated components representing random variations due to fading, shadowing, and noise. When the correlation coefficient approaches 1, there is a strong spatial correlation between the eavesdropping and legitimate channels, while the spatial correlation of channels is weak when close to 0.

For our design, the CSI mismatch is taken into consideration in the joint optimization problem. Since both Bob and Eve are assumed to be regular users of the system as mentioned above, in a TDD-based transmission frame structure [43], Alice can estimate the CSI of Bob (BCSI) and the CSI of Eve (ECSI) through the transmission of pilot sequences with a linear estimator, e.g., by using a minimum mean square error (MMSE) method [54], [55]. However, the estimated BCSI and ECSI will become outdated during transmission due to the time-varying characteristics of the channel and the large number of links that need to be estimated in a multiple RISs deployed system. From the previous work, RIS mainly focuses on scenarios characterized

by low coverage and directed eavesdropping to support low-mobility users in its neighborhood, which predominantly involves slow-fading channel conditions [40], [45], [49], [56]. Following the work in [30], [40], [43], we consider the SR maximization problem under a statistical CSI error model. Through the estimation, we assume the channels expressed as

$$\mathbf{G} = \tilde{\mathbf{G}} + \Delta\mathbf{G}, \mathbf{h} = \tilde{\mathbf{h}} + \Delta\mathbf{h}, \mathbf{g} = \tilde{\mathbf{g}} + \Delta\mathbf{g}, \mathbf{h}_{di} = \tilde{\mathbf{h}}_{di} + \Delta\mathbf{h}_{di}, \quad (11)$$

where $\tilde{\mathbf{G}}$, $\tilde{\mathbf{h}}$, $\tilde{\mathbf{g}}$ and $\tilde{\mathbf{h}}_{di}$ donate the estimated channels, $i = \{u, e\}$, while $\Delta\mathbf{G}$, $\Delta\mathbf{h}$, $\Delta\mathbf{g}$ and $\Delta\mathbf{h}_{di}$ are the estimation error parts, which follow zero-mean complex Gaussian distribution. The quality of the channel estimation is indicated by $\epsilon_G = \mathbb{E}\left[||\Delta\mathbf{G}||_2^2|\tilde{\mathbf{G}}\right]$, $\epsilon_h = \mathbb{E}\left[||\Delta\mathbf{h}||_2^2|\tilde{\mathbf{h}}\right]$, $\epsilon_g = \mathbb{E}\left[||\Delta\mathbf{g}||_2^2|\tilde{\mathbf{g}}\right]$ and $\epsilon_{di} = \mathbb{E}\left[||\Delta\mathbf{h}_{di}||_2^2|\tilde{\mathbf{h}}_{di}\right]$ as considered in [30], [43], [55]. Then, to simplify the description of the uncertain channel model, we follow the methods described in [43] and define a sample space of the channels as $\mathbf{\Omega} \triangleq \{\mathbf{G}_l(\delta), \mathbf{h}_l(\delta), \mathbf{g}_l(\delta), \mathbf{h}_{di}(\delta), \forall\delta, \forall l \in \mathcal{L}\}$, based on the description of imperfect CSI and statistic CSI errors. δ denotes the index of the random realizations of $\mathbf{G}$, $\mathbf{h}$, $\mathbf{g}$, and $\mathbf{h}_{di}$ drawn from the sample space $\mathbf{\Omega}$. Specially, these realizations keep constants for varied δ when under perfect CSI setting.

2.3 Problem Formulation

In this work, we aim to maximize the achievable SR, $R_s(\mathbf{w}, \mathbf{\Theta})$ via a joint design of the active beamforming vector $\mathbf{w}$ at Alice and the passive beamforming matrix $\mathbf{\Theta}$ at the RIS. Assuming perfect CSI of Bob and Eve are available at Alice, the optimization problem is formulated by

$$\text{(P1)} \quad \max_{\mathbf{w},\mathbf{\Theta}} \quad R_s(\mathbf{w}, \mathbf{\Theta}), \quad (12a)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{w}\mathbf{w}^H) \le P, \quad (12b)$$

$$\mathbf{\Theta} \in \mathcal{M}, \quad (12c)$$

where (12b) is the transmitted power constraint, and (12c) is the unit-modulus constraint on each element, i.e., $\mathcal{M} = \left\{\mathbf{\Theta} \middle| \left|\theta_{m,l}\right| = 1, \forall m = 1, \ldots, M,\ l \in \mathcal{L}\right\}$.

Next, we extend our design to a practical scenario with CSI error on both the BCSI and ECSI. Based on the channel model under imperfect CSI, we can see that the SR maximization problem is highly dependent on the random system state. From the channel model discussed above and the properties of transmission frame in RIS-aided system described in [43], [56], we assume that $\mathbf{w}$ is adaptive to the real-time CSI since the channel associate with the active beamforming $\left(\mathbf{h}_l^H\mathbf{\Theta}_l\mathbf{G}_l + \mathbf{h}_{du}^H\right) \in \mathbb{C}^{1\times N}$ with $N \ll M, l \in \mathcal{L}$, is of low-dimension. While the channel associated with the passive beamforming design $\mathbf{G}_l\,\text{diag}\,(\mathbf{h}_l) \in \mathbb{C}^{N\times M}, l \in \mathcal{L}$ is high-dimensional, which is assumed to remain approximately constant within the transmission frame and can be adjusted to the statistics of the random states. This assumption is also valid as RIS is generally used to support low-mobility users in its neighborhood [40], [45], [49], [56]. In other words, we can optimize the active beamforming and passive beamforming at different levels. The expectation of the achievable SR over the channel statistics from the sample space $\mathbf{\Omega}$ is formulated as

$$\text{(P2)} \quad \max_{\mathbf{\Theta},\{\mathbf{w}(\delta),\forall\delta\}} \quad \mathbb{E}_\delta\left[R_s\left(\mathbf{\Theta}, \mathbf{w}(\delta); \delta\right)\right], \quad (13a)$$

$$\text{s.t.} \quad \text{tr}\left(\mathbf{w}(\delta)\mathbf{w}(\delta)^H\right) \le P, \quad \forall\delta, \quad (13b)$$

$$\boldsymbol{\theta} \in \mathcal{M}. \quad (13c)$$

The problem (P2) is intractable due to the uncertain system state in the coupled variables. To tackle this problem, we decompose problem (P2) into two simplified subproblems as [57], which can be solved alternately. Based on the solution of the first subproblem associated with the current random state, the second subproblem is optimized recursively to obtain the maximum achievable SR, which results in a two-level stochastic optimization procedure as elaborated explicitly later in Sect. 4.

3. The LCFP Algorithm for Secrecy Rate Optimization

In this section, we aim to solve the SR maximization problem by taking both the computational complexity and the extension to complex scenarios into consideration.

3.1 Problem Transformation

We note that (P1) is challenging to solve, where the objective function is not jointly concave with respect to both the coupled variables $(\mathbf{w}, \mathbf{\Theta})$, and it contains the unit-modulus constraint (12c). To address the non-convex SR maximization problem with the assumption that the BCSI and ECSI are available at Alice, we aim to develop a low-complexity PLS enhancement algorithm based on the fractional programming (FP) technique, the prox-linear BCD updating technique, and the SCA method. For clarity, a summary of the algorithm, referred to as the LCFP algorithm, is given in Algorithm 1.

To deal with the logarithms and fractions parts of the

Algorithm 1 The Low-complexity Fractional Programming (LCFP) Algorithm.

1: **Input:** $t = 0$, given $\mathbf{w}^0, \boldsymbol{\theta}^0, \tau^0$.
2: **repeat**
3: Set $t = t + 1$;
4: Update the value of $a_1^{(t)}, a_2^{(t)}, \rho^{(t)}$ by (20a) (20b) (20c), and $\nabla R_s, \hat{\lambda}$ by (24) (26);
5: Obtain the optimal solution $\mathbf{w}^{(t)}$ by (25);
6: Construct $\hat{\boldsymbol{\theta}}^{(t)}, \mathbf{v}_u$ and $\mathbf{v}_e$ based on given $\mathbf{w}^{(t)}$ and $\boldsymbol{\theta}^{(t-1)}$;
6: Transform $\hat{\boldsymbol{\theta}}^{(t)}$ to $\hat{\boldsymbol{\phi}}^{(t)}$;
7: Update the $\mathbf{U}$ and $\mathbf{z}$ in (32) and (33);
8: Design step size τ using (36);
9: Update $\hat{\boldsymbol{\phi}}^{(t)}$ with (38), and transform $\hat{\boldsymbol{\phi}}^{(t)}$ to $\hat{\boldsymbol{\theta}}^{(t)}$;
10: With given $\hat{\boldsymbol{\theta}}^{(t)}$, update $\boldsymbol{\theta}^{(t)}$;
11: **Until** the stopping criterion is met.
12: **Output:** $\mathbf{w}, \boldsymbol{\theta}$.

objective function, we first reformulate the objective function by letting $R_s = R_U + R_E$, where R_U and R_E are expressed as

$$\begin{aligned} R_U &= \log\left(1 + \frac{\left\|\left(\sum_{l=1}^{L} \mathbf{h}_l^H \boldsymbol{\Theta}_l \mathbf{G}_l + \mathbf{h}_{du}^H\right)\mathbf{w}\right\|^2}{\sigma^2}\right), \\ R_E &= -\log\left(1 + \frac{\left\|\left(\sum_{l=1}^{L} \mathbf{g}_l^H \boldsymbol{\Theta}_l \mathbf{G}_l + \mathbf{h}_{de}^H\right)\mathbf{w}\right\|^2}{\sigma_e^2}\right). \end{aligned} \tag{14}$$

To construct a more tractable surrogate function to approximate the original non-convex objective function, with $\bar{\mathbf{h}}_{u,l} = \operatorname{diag}\left(\mathbf{h}_l^H\right)\mathbf{G}_l$ and $\bar{\mathbf{h}}_{e,l} = \operatorname{diag}\left(\mathbf{g}_l^H\right)\mathbf{G}_l$ we introduce auxiliary variables $\gamma_{u,1}$, $\gamma_{u,2}$, $\gamma_{e,2}$, and $\gamma_{e,p}$ defined as

$$\gamma_{u,1} = \left\|\left(\sum_{l=1}^{L} \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{u,l} + \mathbf{h}_{du}^H\right)\mathbf{w}\right\|^2, \tag{15a}$$

$$\gamma_{u,2} = \left\|\left(\sum_{l=1}^{L} \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{u,l} + \mathbf{h}_{du}^H\right)\mathbf{w}\right\|^2 + \sigma^2, \tag{15b}$$

$$\gamma_{e,2} = 1 + \frac{\left\|\left(\sum_{l=1}^{L} \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{e,l} + \mathbf{h}_{de}^H\right)\mathbf{w}\right\|^2}{\sigma_e^2}, \tag{15c}$$

$$\gamma_{e,p} = 1 + \frac{\left\|\sum_{l=1}^{L} \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{e,l} + \mathbf{h}_{de}^H\right\|_F^2 P}{\sigma_e^2}. \tag{15d}$$

Next, by taking advantage of the Lagrangian dual transform technique as [58], R_U can be reformulated by

$$R_U = \log(1 + a_1) - a_1 + (1 + a_1)\frac{\gamma_{u,1}}{\gamma_{u,2}}, \tag{16}$$

where a_1 is the auxiliary variable introduced by the transformation. We note that R_E contains a negative logarithm function and is thus difficult to solve directly via the Lagrangian dual transformation. Then, we introduce the auxiliary variables $\gamma_{e,p}$ and $\gamma_{e,2}$ to construct an equivalent expression of R_E, which is given by

$$R_E = \log\left(1 + \frac{\gamma_{e,p} - \gamma_{e,2}}{\gamma_{e,2}}\right) - \log(\gamma_{e,p}). \tag{17}$$

From the Cauchy-Schwarz inequality, it holds $\gamma_{e,2} = \left\|\left(\sum_{l=1}^{L} \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{e,l} + \mathbf{h}_{de}^H\right)\mathbf{w}\right\|^2 \le \left(\left\|\sum_{l=1}^{L} \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{e,l} + \mathbf{h}_{de}^H\right\|_F^2 \|\mathbf{w}\|^2\right) \le \left(\left\|\sum_{l=1}^{L} \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{e,l} + \mathbf{h}_{de}^H\right\|_F^2 P\right) = \gamma_{e,p}$, and thus the fractional terms in logarithm is non-negative. Then, by applying the Lagrangian dual transform technique again, R_E can be rewritten in the following form

$$R_E = \log(1 + a_2) - a_2 + (1 + a_2)\frac{\gamma_{e,p} - \gamma_{e,2}}{\gamma_{e,p}} - \log(\gamma_{e,p}), \tag{18}$$

where a_2 is the auxiliary variable. Noting that the objective is still non-convex, we transform the original objective into the problem (P1.1) based on the idea of decoupled optimization of numerators and denominators in the quadratic transform technique [58].

$$\begin{aligned} (\text{P1.1}) \quad & \max_{\mathbf{w},\boldsymbol{\theta},\mathrm{a}_1,\mathrm{a}_2,\rho} \tilde{\mathrm{R}}_\mathrm{s} = \log(1 + \mathrm{a}_1) + \log(1 + \mathrm{a}_2) \\ & - \log(1 + \frac{\left\|\sum_{l=1}^{L} \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{e,l} + \mathbf{h}_{de}^H\right\|_F^2 P}{\sigma_e^2}) \\ & + 2\rho\sqrt{(1 + a_1)}\Re\left\{\left(\sum_{l=1}^{L} \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{u,l} + \mathbf{h}_{du}^H\right)\mathbf{w}\right\} \\ & - \rho^2\left(\left\|\left(\sum_{l=1}^{L} \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{u,l} + \mathbf{h}_{du}^H\right)\mathbf{w}\right\|^2 + \sigma^2\right) - (a_1 + a_2) + (1 + a_2) \\ & \times \frac{\left\|\sum_{l=1}^{L} \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{e,l} + \mathbf{h}_{de}^H\right\|_F^2 P - \left\|\left(\sum_{l=1}^{L} \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{e,l} + \mathbf{h}_{de}^H\right)\mathbf{w}\right\|^2}{\sigma_e^2 + \left\|\sum_{l=1}^{L} \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{e,l} + \mathbf{h}_{de}^H\right\|_F^2 P}, \\ & \text{s.t.} \qquad (12b), (12c), \end{aligned} \tag{19}$$

where ρ is the new auxiliary introduced by transformation.

3.2 Active Beamforming Design

We fix the passive beamforming variable $\boldsymbol{\theta}$. Note that $\tilde{R}_s$ is convex with the $\mathbf{w}$, a_1, a_2, ρ when other variables are fixed. Hence, we can employ the block coordinate descent method [59] to attain the optimal values of a_1, a_2 and ρ by setting the derivative to zero, while keeping other variables fixed. The closed-form solution for updating the variable a_1, a_2 and ρ at the t-th iteration can be given by

$$a_1^{(t)} = \left(\sqrt{1 + \frac{4}{\vartheta_1^{(t-1)}}} + 1\right)\frac{\vartheta_1^{(t-1)}}{2}, \tag{20a}$$

$$a_2^{(t)} = \frac{\gamma_{e,p}^{(t-1)} - \gamma_{e,2}^{(t-1)}}{\gamma_{e,2}^{(t-1)}}, \tag{20b}$$

$$\rho^{(t)} = \frac{\sqrt{1 + a_1^{(t-1)}}\Re\left\{\vartheta_2^{(t-1)}\right\}}{\gamma_{u,2}^{(t-1)}}, \tag{20c}$$

where $\vartheta_1 = \rho^2\gamma_{u,1}$ and $\vartheta_2 = \left(\sum_{l=1}^{L} \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{u,l} + \mathbf{h}_{du}^H\right)\mathbf{w}$. For the variable $\mathbf{w}$, the dual variable $\hat{\lambda}$ is introduced to deal with the power constraint. When other variables are given, by constructing a Lagrangian function of the objective function in the problem (P1.1), the closed-form solution to the active beamforming subproblem is

$$\begin{aligned} \mathbf{w}^{(t)} = {} & \sqrt{1 + a_1^{(t-1)}}\Re\left\{\sum_{l=1}^{L} \bar{\mathbf{b}}_{u,l}^{(t-1)} + \mathbf{h}_{du}\right\}\Bigg[\left(\rho^{(t-1)}\right)^2 \bar{\mathbf{B}}_u^{(t-1)} \\ & + \left(1 + a_2^{(t-1)}\right)\frac{\bar{\mathbf{B}}_e^{(t-1)}}{\left\|\sum_{l=1}^{L} \bar{\mathbf{b}}_{e,l}^{(t-1)} + \mathbf{h}_{de}\right\|_F^2 P + \sigma_e^2} + \hat{\lambda}\mathbf{I}_N\Bigg]^{-1}\rho^{(t-1)}, \end{aligned} \tag{21}$$

where $\bar{\mathbf{B}}_u = \sum_{l=1}^{L} \bar{\mathbf{b}}_{u,l}^H \bar{\mathbf{b}}_{u,l} + \sum_{l=2}^{L}\sum_{i=1}^{l-1} \bar{\mathbf{b}}_{u,l}^H \bar{\mathbf{b}}_{u,i} + 2\mathbf{h}_{du}\bar{\mathbf{b}}_{u,l} +$

$\mathbf{h}_{du}\mathbf{h}_{du}^H$, $\bar{\mathbf{B}}_e = \sum_{l=1}^{L} \bar{\mathbf{b}}_{e,l}^H \bar{\mathbf{b}}_{e,l} + \sum_{l=2}^{L}\sum_{i=1}^{l-1} \bar{\mathbf{b}}_{e,l}^H \bar{\mathbf{b}}_{e,i} + 2\mathbf{h}_{de}\bar{\mathbf{b}}_{e,l} + \mathbf{h}_{de}\mathbf{h}_{de}^H$, $\bar{\mathbf{b}}_{u,l} = \sum_{l=1}^{L} \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{u,l}$ and $\bar{\mathbf{b}}_{e,l} = \boldsymbol{\theta}_l^H \bar{\mathbf{h}}_{e,l}$. To obtain the optimal value of the Lagrangian multiplier, the binary search method is a possible solution which, however, could greatly increase the complexity. Thus, inspired by [43], we use the prox-linear BCD updating technique to reduce the iteration times for searching $\hat{\lambda}$.

Note that the block for $\mathbf{w}$ is a convex problem while all other blocks are fixed at their last updated values, which satisfies the constraint required for convergence in the prox-linear BCD updates [60]. Then, we attain the optimal solution of $\mathbf{w}$ at each iteration by updating

$$\begin{aligned}\mathbf{w}^{(t+1)} =& \arg\min_{\mathbf{w}}\Big(\langle -\nabla\tilde{R}_s(\tilde{\mathbf{w}}^{(t)}), \mathbf{w}-\tilde{\mathbf{w}}^{(t)}\rangle \\ &+ \frac{C^t}{2}\left\|\mathbf{w}-\tilde{\mathbf{w}}^{(t)}\right\|^2\Big), \\ &\text{s.t.} \quad (12b),\end{aligned} \tag{22}$$

where the Lipschitz constant of the block-partial gradient is expressed as

$$C = \left\|\left(2\rho^2\bar{\mathbf{B}}_u + 2(1+a_2)\frac{\bar{\mathbf{B}}_e}{\left\|\sum_{l=1}^{L}\bar{\mathbf{b}}_{e,l}+\mathbf{h}_{de}\right\|_F^2 P+\sigma_e^2}\right)\right\|_F. \tag{23}$$

The extrapolated point $\tilde{\mathbf{w}}^{(t)} = \mathbf{w}^{(t)} + \varsigma\left(\mathbf{w}^{(t)} - \mathbf{w}^{(t-1)}\right)$, where the extrapolation weight $\varsigma^{(t)} = \min(\hat{\varsigma}^{(t)}, \delta\sqrt{(C^{(t-1)})/C^{(t)}})$ satisfies $\delta < 1$ and $\hat{\varsigma}^t = \left(e^{(t-1)}-1\right)/e^{(t)}$ with $e^0 = 1$, and $e^{(t)} = (1+\sqrt{1+4(e^{(t-1)})^2})/2$. From (19), the block-partial gradient at $\tilde{\mathbf{w}}^{(t)}$ can be obtained as

$$\begin{aligned}\nabla\tilde{R}_s(\tilde{\mathbf{w}}^{(t)}) =& 2\rho\sqrt{1+a_1}\Re\left\{\sum_{l=1}^{L}\bar{\mathbf{b}}_{u,l}+\mathbf{h}_{du}\right\} - 2\rho^2\bar{\mathbf{B}}_u\tilde{\mathbf{w}}^{(t)} \\ &- 2(1+a_2)\left(\frac{\bar{\mathbf{B}}_e\tilde{\mathbf{w}}^{(t)}}{\left\|\sum_{l=1}^{L}\bar{\mathbf{b}}_{e,l}+\mathbf{h}_{de}\right\|_F^2 P+\sigma_e^2}\right).\end{aligned} \tag{24}$$

Since the problem (P1.1) is convex with respect to $\mathbf{w}$, based on the Lagrange multiplier method, the optimal active beamforming solution can be obtained

$$\mathbf{w}^{(t+1)} = \frac{1}{C^{(t)}-2\hat{\lambda}}\left(C^{(t)}\tilde{\mathbf{w}}^{(t)} + \nabla\tilde{R}_s(\tilde{\mathbf{w}}^{(t)})\right), \tag{25}$$

where the Lagrange multiplier for the power constraint (12b) can be updated by

$$\hat{\lambda}^* = \frac{C^{(t)}}{2} - \frac{\left\|C^{(t)}\tilde{\mathbf{w}}^{(t)} + \nabla R_s(\tilde{\mathbf{w}}^{(t)})\right\|^2}{2P}. \tag{26}$$

By reducing the iteration times for searching λ, and utilizing a more efficient technique than the original BCD technique as proved in [60], the complexity of the algorithm for active beamforming optimization can be greatly reduced, which will be further discussed in Sect. 5.

3.3 Passive Beamforming Design

After updating the value of $\mathbf{w}$ according to (25), we aim to optimize the passive beamforming variable $\boldsymbol{\theta}$. Unfortunately, the subproblem of passive beamforming optimization is non-convex and more difficult than directly optimizing the active beamforming due to the unit-modulus constraint. To address this difficulty, we first reformulate the expressions of the objective function R_S in terms of $\boldsymbol{\theta}$ by constructing $\hat{\boldsymbol{\Theta}} = [\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_L]$, $\mathbf{V}_i = [\mathbf{v}_{i,1}, \mathbf{v}_{i,2}, \cdots, \mathbf{v}_{i,L}]$, $i \in \{u, e\}$, where $\mathbf{v}_{u,l} = \mathrm{diag}(\mathbf{h}_l^H)\mathbf{G}_l\mathbf{w}$, $\mathbf{v}_{e,l} = \mathrm{diag}(\mathbf{g}_l^H)\mathbf{G}_l\mathbf{w}$. We also denote $\hat{\boldsymbol{\theta}} = \mathrm{vec}(\hat{\boldsymbol{\Theta}})$, $\mathbf{v}_i = \mathrm{vec}(\mathbf{V}_i)$, and

$$\bar{\gamma}_{u,1} = \left\|\mathbf{h}_{du}^H\mathbf{w} + \hat{\boldsymbol{\theta}}^H\mathbf{v}_u\right\|^2, \tag{27a}$$

$$\bar{\gamma}_{u,2} = \left\|\mathbf{h}_{du}^H\mathbf{w} + \hat{\boldsymbol{\theta}}^H\mathbf{v}_u\right\|^2 + \sigma^2, \tag{27b}$$

$$\bar{\gamma}_{e,2} = 1 + \frac{\left\|\mathbf{h}_{de}^H\mathbf{w} + \hat{\boldsymbol{\theta}}^H\mathbf{v}_e\right\|^2}{\sigma_e^2}, \tag{27c}$$

$$\bar{\gamma}_{e,1} = 1 + \frac{\left(\left\|\mathbf{h}_{de}^H\mathbf{w}\right\| + \sqrt{M}\left\|\mathrm{tr}(\mathbf{V}_e)\right\|\right)^2}{\sigma_e^2}. \tag{27d}$$

Then, R_U in (14) is equivalent to

$$R_U = \log(1+a_1) + (1+a_1)\frac{\bar{\gamma}_{u,1}}{\bar{\gamma}_{u,2}} - a_1. \tag{28}$$

Based on the triangle inequality and the constraint (12c), we can obtain $\|\mathbf{h}_{de}^H\mathbf{w}+\hat{\boldsymbol{\theta}}^H\mathbf{v}_e\|^2 \le (\left\|\mathbf{h}_{de}^H\mathbf{w}\right\| + \|\hat{\boldsymbol{\theta}}^H\mathbf{v}_e\|)^2 \le (\left\|\mathbf{h}_{de}^H\mathbf{w}\right\| + \sqrt{M}\left\|\mathrm{tr}(\mathbf{V}_e)\right\|)^2$, which shows that the function $(\bar{\gamma}_{e,1}-\bar{\gamma}_{e,2})$ is non-negative. Again, by constructing an equivalent expression to R_E and applying the Lagrangian dual transform technique, we can obtain

$$\begin{aligned}R_E =& \log(1+a_2) - \log(\bar{\gamma}_{e,1}) \\ &+ (1+a_2)\left(\frac{\bar{\gamma}_{e,1}-\bar{\gamma}_{e,2}}{\bar{\gamma}_{e,1}}\right) - a_2.\end{aligned} \tag{29}$$

After dropping the constant terms and applying the quadratic transform technique, the original optimization problem can be reformulated into (P1.2), where $D_i = \mathbf{h}_{di}^H\mathbf{w}, i \in \{u, e\}$.

$$\begin{aligned}(\text{P1.2}) \quad &\max_{\boldsymbol{\theta},\mathrm{a}_1,\mathrm{a}_2,\hat{\rho}} \hat{\mathrm{R}}_\mathrm{s} = 2\hat{\rho}\sqrt{(1+a_1)}\Re\left\{\mathrm{D_u} + \hat{\boldsymbol{\theta}}^\mathrm{H}\mathbf{v}_\mathrm{u}\right\} \\ &- \hat{\rho}^2\left(\left\|D_u + \hat{\boldsymbol{\theta}}^H\mathbf{v}_u\right\|^2 + \sigma^2\right) \\ &+ (1+a_2)\left(\frac{(\|D_e\| + \sqrt{M}\left\|\mathrm{tr}(\mathbf{V}_e)\right\|)^2 - \left\|D_e + \hat{\boldsymbol{\theta}}^H\mathbf{v}_e\right\|^2}{\sigma_e^2 + (\|D_e\| + \sqrt{M}\left\|\mathrm{tr}(\mathbf{V}_e)\right\|)^2}\right), \\ &\text{s.t.} \quad (12c).\end{aligned} \tag{30}$$

Note that problem (P1.2) is convex with the auxiliary variable when other variables are fixed. Hence, the auxiliary

variables introduced by transformation techniques at the t-th iteration are given by

$$\hat{\rho}^{(t)} = \frac{\sqrt{(1+a_1^{(t-1)})}\Re\left\{D_u^{(t-1)} + \left(\hat{\boldsymbol{\theta}}^{(t-1)}\right)^H \mathbf{v}_u^{(t-1)}\right\}}{\|D_u^{(t-1)} + (\hat{\boldsymbol{\theta}}^{(t-1)})^H \mathbf{v}_u^{(t-1)}\|^2 + \sigma^2}. \tag{31}$$

It can be found that the problem (P1.2) is non-convex with respect to $\boldsymbol{\theta}$ when fixing all other variables, because of the unit-modulus constraints. To handle this problem, we let the phase shift as a substitute for $\theta_{m,l} = e^{j\phi_{m,l}}$, and construct $\hat{\boldsymbol{\Phi}} = [\boldsymbol{\phi}_1, \cdots, \boldsymbol{\phi}_l, \cdots, \boldsymbol{\phi}_L]$, where $\boldsymbol{\phi}_l = [\phi_{1,l}, \cdots, \phi_{M,l}]^T$. Then, the objective function can be transformed into a quadric form by letting

$$\mathbf{U} = \hat{\rho}^2 \mathbf{v}_u \mathbf{v}_u^H + (1+a_2)\frac{\mathbf{v}_e \mathbf{v}_e^H}{\sigma_e^2 + (\|D_e\| + \sqrt{M}\,\|\mathrm{tr}(\mathbf{V}_e)\|)^2}, \tag{32}$$

$$\mathbf{z} = \hat{\rho}\sqrt{(1+a_1)}\Re\{\mathbf{v}_u\} - \hat{\rho}^2 (D_u \mathbf{v}_u) - (1+a_2)\frac{D_e \mathbf{v}_e}{\sigma_e^2 + (\|D_e\| + \sqrt{M}\,\|\mathrm{tr}(\mathbf{V}_e)\|)^2}. \tag{33}$$

Therefore, the passive beamforming optimization subproblem can be reformulated as

$$\text{(P1.3)} \quad \min_{\hat{\boldsymbol{\phi}}} \quad \left[\left(\mathrm{e}^{\mathrm{j}\hat{\boldsymbol{\phi}}}\right)^{\mathrm{H}} \mathbf{U} \mathrm{e}^{\mathrm{j}\hat{\boldsymbol{\phi}}} - 2\Re\left\{(\mathrm{e}^{\mathrm{j}\hat{\boldsymbol{\phi}}})^{\mathrm{H}} \mathbf{z}\right\}\right], \tag{34}$$

where $\hat{\boldsymbol{\phi}} = \mathrm{vec}(\hat{\boldsymbol{\Phi}})$. To solve the non-convex problem in (34), we apply an inexact BCD method called the block successive convex approximation (BSCA) method [61] to guarantee convergence. Following the updating rule, we construct a convex second-order approximation of the objective function $\hat{R}_s$ at $\hat{\boldsymbol{\phi}}^{(t)}$ as the surrogate function, which is expressed as

$$\hat{R}_s(\hat{\boldsymbol{\phi}}|\hat{\boldsymbol{\phi}}^{(t)}) = \hat{R}_s\left(\hat{\boldsymbol{\phi}}^{(t)}\right) + \nabla\hat{R}_s\left(\hat{\boldsymbol{\phi}}^{(t)}\right)\left(\hat{\boldsymbol{\phi}} - \hat{\boldsymbol{\phi}}^{(t)}\right) + \frac{1}{2\tau}\left\|\hat{\boldsymbol{\phi}} - \hat{\boldsymbol{\phi}}^{(t)}\right\|^2. \tag{35}$$

Since the surrogate function (35) is a local approximation of the original function (34), to search for an optimal value with sufficient decrease, we adopt the line search with the Armijo step size selection rule, where the step size τ is the largest element in $\{\tau^0\beta^r\}_{r=0,1,\cdots}$, $\tau^0 > 0$, $\beta \in (0,1)$, satisfying

$$\hat{R}_s\left(\hat{\boldsymbol{\phi}}^{(t)}\right) - \hat{R}_s(\hat{\boldsymbol{\phi}}) \geq -\varepsilon\tau\left\|\nabla\hat{R}_s\left(\hat{\boldsymbol{\phi}}^{(t)}\right)\right\|^2, \tag{36}$$

where $\varepsilon \in (0,1)$, and the gradient with respect to $\hat{\boldsymbol{\phi}}$ is given by

$$\nabla\hat{R}_s(\hat{\boldsymbol{\phi}}) = 2\Re\left\{-je^{-j\hat{\boldsymbol{\phi}}}\left(\mathbf{U}e^{j\hat{\boldsymbol{\phi}}} - \mathbf{z}\right)\right\}. \tag{37}$$

Based on the minimum value of the surrogate function in (35), the gradient projection updating for the optimization variable $\boldsymbol{\phi}$ is quadratic programming with a closed-form solution

$$\hat{\boldsymbol{\phi}}^{(t)} = \hat{\boldsymbol{\phi}}^{(t-1)} - \tau\nabla\hat{R}_s. \tag{38}$$

With the function (34) continuously differentiable, τ can be properly chosen to make the surrogate function satisfy the following constraint

$$\hat{R}_s\left(\hat{\boldsymbol{\phi}}|\hat{\boldsymbol{\phi}}^{(t)}\right) \geq \hat{R}_s(\hat{\boldsymbol{\phi}}), \tag{39}$$

which guarantees the convergence of the algorithm based on the proposition in [61]. The equality can be obtained, when $\hat{\boldsymbol{\phi}} = \hat{\boldsymbol{\phi}}^{(t)}$.

The LCFP algorithm also can be easily extended to a RIS-aided SR maximization optimization problem associated with CSI errors, since the coordinated optimization algorithms for subproblems satisfy the convergence conditions to solve a two-level stochastic non-convex optimization problem [62]. Therefore, considering reducing the complexity and the cost of obtaining CSI, we extend the LCFP algorithm to a more practical scenario in the next section.

4. The LCRFP Algorithm for Secrecy Rate Optimization with Imperfect CSI

In this section, we investigate the SR maximization problem under a more practical scenario where both the CSI of Bob and Eve are imperfect.

4.1 Active Beamforming Design at the First Level

As analyzed in Sect. 2, problem (P2) is a two-level stochastic non-convex optimization problem containing expectation operators with respect to the random system state, which is more difficult to solve compared with (P1). To address this challenging problem, we propose a robust and low-complexity PLS enhancement algorithm based on the statistical CSI error model to maximize the SR. Our method can reduce the impact of the CSI mismatch by combining the benefits of the stochastic non-convex optimization techniques [62] and the LCFP algorithm, which is referred to as the LCRFP algorithm, as presented in Algorithm 2. This algorithm consists of three major steps. First, based on the problem transformation technique in the LCFP algorithm, the problem (P2) is decoupled into two subproblems, which can be optimized at different levels. Next, in the case of given $\bar{\boldsymbol{\theta}}$, we have the first-level subproblem with respect to $\mathbf{w}(\delta)$, which is the first-level optimization variable under the system state. To deal with the random system state and satisfy the assumptions when employing the stochastic successive convex approximation techniques [62], we initially transform the objective into a minimization problem. Then, following a similar approach used in the LCFP algorithm (Sect. 3.1) for active beamforming design, we employ the FP transformation technique to address logarithmic and fractional components by introducing auxiliary variables $\bar{a}_1$, $\bar{a}_2$ and $\bar{\rho}$. The first-level optimization subproblem can be reformulated as

Algorithm 2 The LCRFP Algorithm for Imperfect CSI.

Input: $\{\varrho^{(t)}\}$, initialize $t = 0$, $\mathbf{w}(\delta)^0$, δ^0, $\boldsymbol{\theta}^0$, and τ^0.
  repeat
    Set $t = t + 1$, and realize the channel about $\delta^{(t)}$;
    Solve the first-level subproblem $(P1.1)$ to obtain the optimal solution $\mathbf{w}(\delta)^{(t)}$ with the given $\delta^{(t)}$ and auxiliary variables;
    Construct $\hat{\boldsymbol{\theta}}^{(t)}$, $\mathbf{v}_u$ and $\mathbf{v}_e$ based on given $\mathbf{w}(\delta)^{(t)}$ and $\boldsymbol{\theta}^{(t-1)}$;
    Transform $\hat{\boldsymbol{\theta}}^{(t)}$ to $\hat{\boldsymbol{\phi}}^{(t)}$;
    Update the $\mathbf{U}, \mathbf{z}$ based on the solution of the first-level subproblem $\mathbf{w}(\delta)^{(t)}$ and $\delta^{(t)}$;
    Calculate the $\nabla_{\hat{\boldsymbol{\phi}}} k\left(\hat{\boldsymbol{\phi}}, \mathbf{w}(\delta); \delta\right)$ by (45);
    Update the gradient $\mathbf{k}_{\hat{\boldsymbol{\phi}}}^{(t)}$ by (44) based on the solution of $\nabla_{\hat{\boldsymbol{\phi}}} k\left(\hat{\boldsymbol{\phi}}, \mathbf{w}(\delta); \delta\right)$;
    Design τ using Armijo rule according to $\hat{\boldsymbol{\phi}}^{(t)}$, $\mathbf{w}(\delta)^{(t)}$;
    Update $\hat{\boldsymbol{\phi}}^{(t)}$ with (47), and transform $\hat{\boldsymbol{\phi}}^{(t)}$ to $\hat{\boldsymbol{\theta}}^{(t)}$;
    With given $\hat{\boldsymbol{\theta}}^{(t)}$, update $\boldsymbol{\theta}^{(t)}$;
  Until the stopping criterion is met.
Output: $\mathbf{w}, \boldsymbol{\theta}$.

$$(\text{P2.1}) \quad \min_{\mathbf{w}(\delta), \bar{a}_1, \bar{a}_2, \bar{\rho}} \; -\tilde{R}_s\left(\bar{\boldsymbol{\theta}}, \mathbf{w}(\delta), \bar{a}_1, \bar{a}_2, \bar{\rho}; \delta\right), \tag{40a}$$

$$\text{s.t.} \quad \operatorname{tr}\left(\mathbf{w}(\delta)\,\mathbf{w}(\delta)^H\right) \leqslant P, \quad \forall \delta, \tag{40b}$$

where $\tilde{R}_s$ is the same form as in (P1.1). As problem (P2.1) is convex with the $\bar{a}_1, \bar{a}_2, \bar{\rho}$ when other variables are fixed, we can obtain a similar closed-form solution of auxiliary variables $\bar{a}_1$, $\bar{a}_2$ and $\bar{\rho}$ as in (20a) (20b) (20c). To derive a closed-form solution to $\mathbf{w}(\delta)$, we first construct a Lagrangian function of the objective to deal with the power constraint. Then, by employing the prox-linear BCD updating technique, which is the same as Algorithm 1 and omitted for simplicity, we obtain the optimal active beamforming solution as depicted in (25). Therefore, during the t-th iteration, after realizing the channel about $\delta(t)$, the auxiliary variables and the optimal solution $\mathbf{w}(\delta)$ can be updated, respectively.

To obtain a stationary solution for the first-level subproblem, we analyze the convergence of the algorithm devised for handling this subproblem. It has been established in [62] that the algorithm design must adhere to three conditions to guarantee convergence. The first condition is met by properly selecting the initial point. Subsequently, the updating equation for each variable within $\mathbf{w}(\delta), a_1, a_2$ exhibits Lipschitz continuity when the other two variables are held constant, satisfying the second convergence criterion. Lastly, we fulfill the third convergence requirement by leveraging the verified global convergence of the prox-linear block coordinate method as outlined in [60]. Therefore, by adhering to these conditions, the optimization algorithm for the first-level subproblem is capable of converging, leading to the stationary solution $\mathbf{w}(\delta)^*$.

4.2 Passive Beamforming Design at the Second Level

After obtaining the corresponding first-level variable $\mathbf{w}(\delta)^{(t)}$ with respect to a new realization of the channel realizations about $\delta^{(t)}$ by running the first-level algorithm in the t-th iteration, we have the second-level subproblem $(P2.2)$ as followed

$$(\text{P2.2}) \quad \min_{\hat{\boldsymbol{\phi}}} \quad k(\hat{\boldsymbol{\phi}}) = \mathbb{E}_{\delta}\left[-\tilde{R}_s\left(\hat{\boldsymbol{\phi}}, \mathbf{w}(\delta); \delta\right)\right], \tag{41}$$

where $\hat{\boldsymbol{\phi}}$ is the second-level optimization variable, which is a substitute for $\hat{\boldsymbol{\theta}}^{(t)}$ to deal with the non-convex constraint, which is constructed in the method of Sect. 3.3. As in the passive beamforming design of the LCFP algorithm, the surrogate function $\bar{k}(\hat{\boldsymbol{\phi}})$ is designed as a convex approximation of the objective function of the problem $(P2.2)$. Therefore, the optimal solution $\hat{\boldsymbol{\phi}}^{(t)}$ is solved as

$$\hat{\boldsymbol{\phi}}^{(t)} = \arg\min_{\hat{\boldsymbol{\phi}}} \quad \bar{k}^{(t)}(\hat{\boldsymbol{\phi}}), \tag{42}$$

To obtain a stationary solution to the second-level subproblem, we first design a convex surrogate function $\bar{k}(\hat{\boldsymbol{\phi}})$ based on the method in (35) of the LCFP algorithm, and this surrogate function satisfy the assumptions of the stochastic non-convex optimization to ensure the convergence of the algorithm, as shown in (39). The approximation of the objective function is expressed as

$$\bar{k}^{(t)}(\hat{\boldsymbol{\phi}}) = k(\hat{\boldsymbol{\phi}})^{(t)} + \mathbf{k}_{\hat{\boldsymbol{\phi}}}^{(t)}\left(\hat{\boldsymbol{\phi}} - \hat{\boldsymbol{\phi}}^{(t)}\right) + \tau\left\|\hat{\boldsymbol{\phi}} - \hat{\boldsymbol{\phi}}^{(t)}\right\|^2. \tag{43}$$

Then, we construct the recursive approximation of the second-level subproblem to adapt to the properties of the first-level subproblem solution with the help of the key theorem for surrogate function design in the stochastic non-convex optimization [62]. Therefore, with the initial value $\mathbf{k}_{\hat{\boldsymbol{\phi}}}^{-1} = 0$, the approximation of the partial derivative $\nabla_{\hat{\boldsymbol{\phi}}} k\left(\hat{\boldsymbol{\phi}}^{'}\mathbf{w}(\delta); \delta\right)$, $\mathbf{k}_{\hat{\boldsymbol{\phi}}}$ in (43) updates recursively as

$$\mathbf{k}_{\hat{\boldsymbol{\phi}}}^{(t)} = (1 - \varrho^{(t)})\mathbf{k}_{\hat{\boldsymbol{\phi}}}^{(t-1)} + \varrho^{(t)}\nabla_{\hat{\boldsymbol{\phi}}} k\left(\hat{\boldsymbol{\phi}}^{(t)}, \mathbf{w}(\delta)^{(t)}; \delta^{(t)}\right), \tag{44}$$

where $\varrho^{(t)} \in (0, 1]$ is a sequence satisfying $\sum_t \varrho^{(t)} = \infty$, $\sum_t \left(\varrho^{(t)}\right)^2 < \infty$. And the gradient of $k(\hat{\boldsymbol{\phi}}, \mathbf{w}(\delta); \delta)$ with respect to $\hat{\boldsymbol{\phi}}$ in (44) is given by

$$\nabla_{\hat{\boldsymbol{\phi}}} k\left(\hat{\boldsymbol{\phi}}, \mathbf{w}(\delta); \delta\right) = 2\Re\left\{-je^{-j\hat{\boldsymbol{\phi}}}\left(\mathbf{U}e^{j\hat{\boldsymbol{\phi}}} - \mathbf{z}\right)\right\}. \tag{45}$$

As such, the constant $k(\hat{\boldsymbol{\phi}})^{(t)}$, with the initial value $k_{\hat{\boldsymbol{\phi}}}^{-1} = 0$, can be calculated via a recursive formula as

$$k(\hat{\boldsymbol{\phi}})^{(t)} = (1 - \varrho^{(t)})k(\hat{\boldsymbol{\phi}})^{(t-1)} + \varrho^{(t)} k\left(\hat{\boldsymbol{\phi}}^{(t)}, \mathbf{w}(\delta)^{(t)}; \delta^{(t)}\right). \tag{46}$$

With the expression of $\mathbf{U}$ and $\mathbf{z}$ already given in the previous section, the approximation of the problem $(P2.2)$ is quadratic programming with respect to the passive beamforming variable $\hat{\boldsymbol{\phi}}$, which leads to a closed-form solution given by

$$\hat{\boldsymbol{\phi}}^{(t+1)} = \hat{\boldsymbol{\phi}}^{(t)} - \frac{\mathbf{k}_{\hat{\boldsymbol{\phi}}}^{(t)}}{2\tau}. \tag{47}$$

As described above, with the convergence conditions satisfied in each level of subproblem optimization, the LCRFP algorithm can obtain a stationary solution by solving the non-convex stochastic joint optimization problem (P2) under the impact of imperfect BCSI and ECSI in a RIS-aided mmWave system.

5. Complexity Analysis

In this section, we analyze the complexity of the proposed algorithms and compare them with some widely used methods for SR maximization in RIS-aided systems. First, in the LCFP algorithm, the complexity to update auxiliary variables and the active beamforming variable $\mathbf{w}$ is $O(2LMN + N^2)$, and the updating loop for this part is $I_{\mathbf{w}} = 1$, since the closed-form solutions can be obtained. The complexity of solving (32) and (33) in the passive beamforming optimization is at most $O(L^2M^2)$, and I_A is the updating loop for step searching due to applying the Armijo-type line search. Therefore, the entire complexity of the LCFP algorithm is $O(I_L(I_A(2MN + N^2) + (LM)^2))$, where I_L is the integral updating loop required for the optimization algorithm convergence. Then, we consider the complexity of the PLS enhancement algorithm based on other existing methods. In an SCA-based [42] PLS enhancement algorithm, the complexity of solving the passive beamforming subproblem is $O((LM)^{3.5})$. The MO method can address the unit-modulus constraint in the RIS-aided system, and the complexity of passive beamforming optimization based on the MO method [43] in the PLS enhancement algorithm is $O(I_{M,\boldsymbol{\theta}}L^2M^2)$. At the same time, the SCA technique is used to optimize the active beamforming as in [42]. We also analyze the complexity of the SDR-based [63] PLS enhancement algorithm. Table 1 summarizes the complexity comparison of the proposed algorithms and the aforementioned algorithms.

Evidently, the proposed algorithm has lower computation complexity than other algorithms. Since $N \ll M$, the main computational complexity of schemes lies in the passive beamforming optimization part. Therefore, the complexity of the SCA-based algorithm and the SDR-based algorithm is higher than the LCFP algorithm. Although the similar computational complexity in the passive beamforming optimization part $O((LM)^2)$, the active beamforming optimization part in the LCFP algorithm $O(N^2)$ still has a slightly higher complexity than the MO-based algorithm $O(N^3)$. The comparison shows that the proposed algorithm has lower computation complexity than benchmark algorithms. The complexity of the SCA-based algorithm and the SDR-based algorithm increases faster than the LCFP algorithm with the number of reflecting elements M increasing. Although there is similar computational complexity in the passive beamforming optimization part, the active beamforming optimization part in the LCFP algorithm $O(N^2)$ still has a slightly higher complexity than the MO-based algorithm $O(N^3)$. Moreover, with the variable $\mathbf{w}, \boldsymbol{\theta}$ coupled, the MO-based algorithm optimizes the subproblems independently, which is hard to extend to the imperfect CSI case. In summary, the LCFP algorithm has advantages in extending to some complex scenarios and low complexity, especially when M is relatively large, conforming to the practical communication system.

Table 1 Comparison of algorithm complexity.

Algorithm	Complexity
LCFP algorithm	$O(I_L(I_A(2LMN + N^2) + (LM)^2))$
SCA-based algorithm	$O(I_C(I_{C,\mathbf{w}}N^3 + I_{C,\boldsymbol{\theta}}(LM)^{3.5}))$
MO-based algorithm	$O(I_M(I_{M,\mathbf{w}}N^3 + I_{M,\boldsymbol{\theta}}(LM)^2))$
SDR-based algorithm	$O(I_D(I_{D,\mathbf{w}}N^{6.5} + I_{D,\boldsymbol{\theta}}(LM)^{6.5}))$

Next, we analyze the complexity of the LCRFP algorithm. Besides the advantage of the low complexity of each algorithm to optimize the active beamforming or passive beamforming at different levels due to the basis of the LCFP algorithm, the updating method utilized in this algorithm also brings benefits in reducing the complexity. Compared with other updating methods to solve a two-level stochastic optimization problem, i.e., batch alternating optimization algorithm, which needs to minimize a sample average approximation function as $k_{\mathrm{SAA}}^{(t)}(\boldsymbol{\phi}) \triangleq \sum_{r=1}^{t} k(\boldsymbol{\phi}, \mathbf{w}(\phi^{(t)}, \delta^r)^*, \delta^r)$, where t is the iteration times, the updating method we applied needs fewer iterations. Since it only needs to solve the first-level problem with respect to the current system state $\delta^{(t)}$, batch alternating optimization needs to solve t times associated with all the previous system states to obtain the average approximation. The total complexity of the LCRFP algorithm is $O(I_R(I_AI_f(2LMN + N^2) + (LM)^2))$, where I_f denotes the iterative numbers of the first-level optimization subproblem to find a stationary solution.

6. Numerical Results

In this section, we evaluate the effectiveness of our proposed algorithms via simulation to verify the utility of the PLS enhancement in a RIS-aided mmWave system when jointly considering the non-neglectable practical issues. The schematic system model for the simulated RIS-aided mmWave system is shown in Fig. 1, where both Bob and Eve are regular users of the system. Alice serves as the transmitter equipped with $N = 4$ antennas and is situated at (0 m, 0 m). The location of the legitimate receiver (Bob) is (D_1 m, 0 m). Eve moves along a circular path with a radius of R m around Bob. Here, R, the distance between Bob and Eve, is initially set as 8 m, and the receiving distance denoted by D_1 is predefined as 70 m. To enhance the communication secrecy between Alice and Bob, L RISs are deployed around Bob. The number of the reflecting elements on each RIS is initially set as $M = 50$. The amplitude reflection coefficient is set as $\eta = 1$ to maximize the reflection strength. Assuming the channel $\mathbf{h}_{d,k}$ follows Rayleigh fading and $\mathbf{G}$, $\mathbf{g}$, and $\mathbf{h}$ follow Rician fading as in [43]. In the cascade channels of Alice-to-RIS, RIS-to-Bob, and RIS-to-

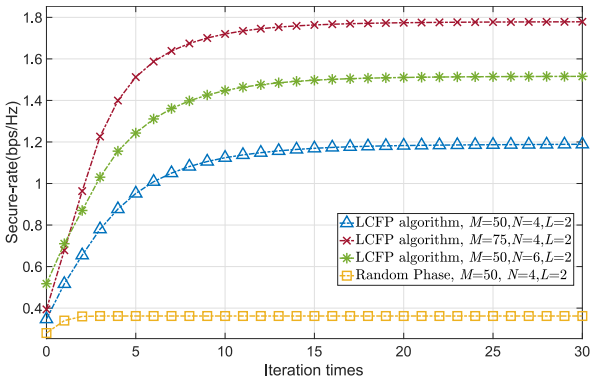


Fig. 2 Convergence of the LCPF algorithm versus different RIS size and transmit antennas.

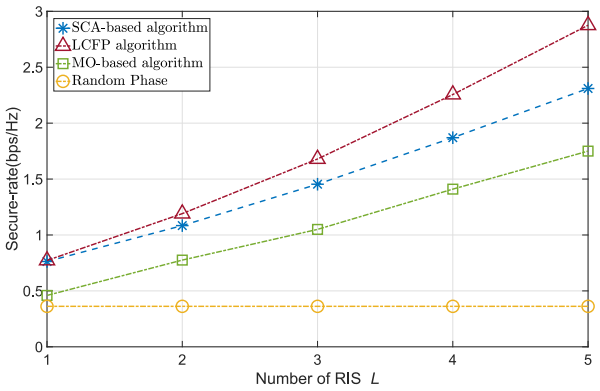


Fig. 3 The algorithms versus different RIS numbers.

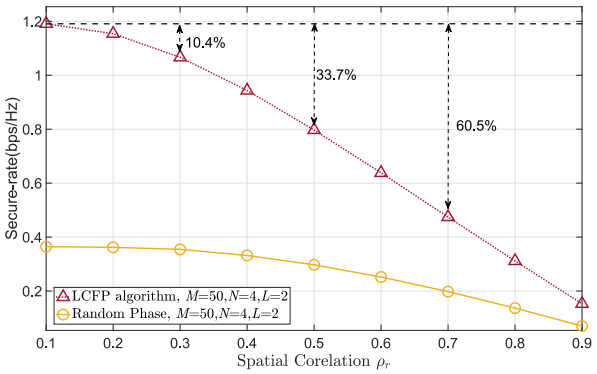


Fig. 4 The algorithms versus different spatial correlation.

Eve, the Rician factor is set as $\kappa = 10$. The path loss exponent of the RIS-related channels is set to 2.2, and that of the direct Alice-Bob and Alice-Eve channels is set to 3.67. The spatial correlation between Bob and Eve is initialized to a low level with $\rho = 0.1$. The total transmit power constraint is set to $P = 5$ dBm.

Firstly, we discuss the convergence behavior of the LCFP algorithm. As illustrated in Fig. 2, the proposed algorithm converges in the limit iterations by assuming the perfect CSI is obtained. Then, we set the random passive beamforming on RIS as a benchmark scheme. It can be seen that the SR of the random phase setting is very low (SR value is 0.36 bps/Hz) due to the more advantageous position of Eve. With a proper passive beamforming optimization of multiple distributed RISs, the proposed algorithm can achieve a secrecy performance gain that outperforms the random phase setting.

Secondly, we show the security performance of the LCFP algorithm versus different numbers of reflecting elements M and antennas N. It can be observed from Fig. 2 that the optimal SR increases with the number of reflecting elements or antennas increasing. Specially, by increasing the reflection number up to 50%, where $M = 75, N = 4$, the achievable SR increases by 49.6%. On the other hand, the achievable SR only increases by 27.5% when increasing the antenna number up to 50%, where $M = 50, N = 6$. Therefore, with the benefits of array gain, increasing the RIS size can provide a more obvious performance improvement than increasing the antenna number.

Next, we illustrate the variation of the achievable SR with the different number of RISs under different RIS-aided PLS enhancement schemes. By increasing the number of the RIS L from 1 to 5 with $P = 5$ dBm, $M = 50, N = 4$, we compare the secrecy performance achieved by the LCFP algorithm with the widely used SCA-based algorithm and MO-based algorithm mentioned in Sect. 5. The SDR-based algorithm is excluded from the simulation due to the high complexity, and the convergence is hard to be guaranteed. By setting the random passive beamforming RIS setting as the lower bound, we treat the MO-based [43] PLS enhancement algorithm and the SCA-based [42] PLS enhancement algorithm as benchmarks. Figure 3 shows that the achievable SR increases as the number of RISs increases under all the algorithms, except the random phase setting scheme. From the results, we can also obtain that the LCFP algorithm can achieve a better secrecy performance compared with the MO-based algorithm. In addition, with the RIS number increasing, the LCFP algorithm can achieve a better secrecy performance than the SCA-based algorithm. Therefore, the proposed LCFP algorithm outperforms the benchmarks in security performance.

Figure 4 evaluates the impact of spatial correlation between the eavesdropping and legitimate channels on security performance. By increasing the spatial correlation coefficient ρ_r from 0.1 (indicating a low level) to 0.9 (indicating a high level), with parameters set as $P = 5$ dBm, $M = 50$, and $N = 4$, we observe a decline in the achievable SR as the spatial correlation grows. Furthermore, this reduction in SR exhibits an amplified trend with increasing spatial correlation. Specifically, compared to a scenario with a low spatial correlation between Bob and Eve ($\rho_r = 0.1$), there is a 10.4% security performance loss as the spatial correlation escalates to 0.3. This reduction becomes more substantial, reaching 33.7%, as the spatial correlation is elevated to 0.5. Notably, an even more pronounced drop of 60.5% is observed with a spatial correlation of 0.7. This is attributed to the reduced channel capacity difference resulting from the decreased spatial correlation between le-

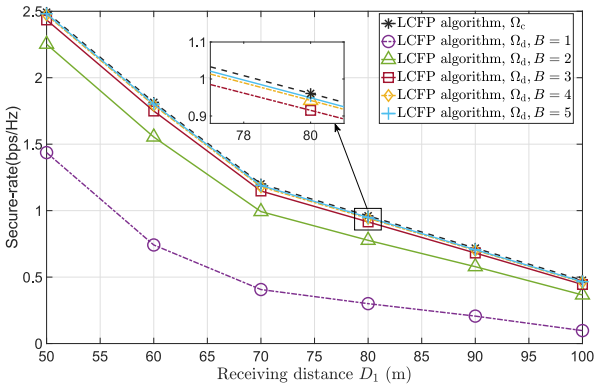


Fig. 5 The performance of proposed algorithm versus different resolution phase shifters.

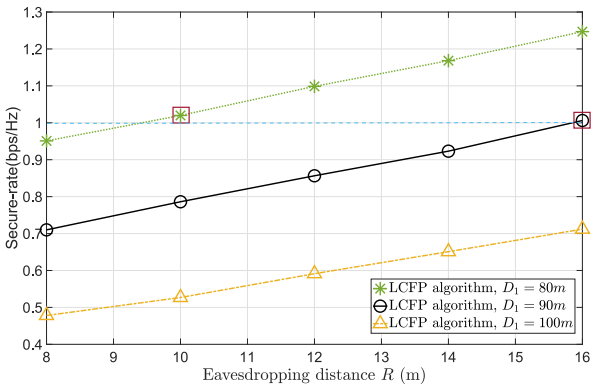


Fig. 6 The performance of proposed algorithm under different spatial positioning.

gitimate and eavesdropping users. However, the proposed LCFP algorithm outperforms the random phase scheme by optimizing the passive beamforming to deal with the performance loss caused by spatial correlation between the eavesdropping and legitimate channels.

Also, we study the impact of the different resolution phase shifters of the RIS on system secrecy performance. Herein, the infinite-resolution-continuous-phase shifter set is defined as $\mathbf{\Omega}_c$ for $\theta_{m,l}$, while the finite-resolution-discrete-phase shifter set as $\mathbf{\Omega}_d$, which are expressed as

$$\mathbf{\Omega}_c = \left\{\theta_{m,l} = e^{j\phi_{m,l}} | \phi_{m,l} \in [0, \cdots, 2\pi)\right\}, \tag{48}$$

$$\mathbf{\Omega}_d = \left\{\theta_{m,l} = e^{j\phi_{m,l}} | \phi_{m,l} = \frac{2\pi n}{2^B}, n \in [0, \cdots, 2^B - 1)\right\}, \tag{49}$$

where B is the number of quantization bits. As shown in Fig. 5, we exploit the secrecy rate gap between finite-resolution phase shifters and infinite-resolution phase shifters. The secrecy rate is plotted versus a range of distances D_1 from 50 m to 100 m between the transmitter (Alice) and receivers (Bob and Eve) with $M = 50, N = 4, L = 2$, and $R = 8$ m. As the phase shifter resolution B increases from 1 to 5, the secrecy rate gap in comparison to infinite-resolution phase shifters diminishes. Specially, at a distance $D_1 = 70$ m, this secrecy rate gap reduces from 0.21 bps/Hz to 0.02 bps/Hz when B varies from 2 to 4. Notably, this secrecy performance enhancement in finite-resolution phase shifters setting exhibits a diminishing trend once B exceeds a value of 3, suggesting a proper threshold for phase resolution. Figure 5 also indicates that increasing the distance D_1 has a detrimental effect on the secrecy performance. Nonetheless, when D_1 remains below a certain threshold, i.e., $D_1 \leq 77$ m, the system is capable of sustaining a secrecy rate above 1.0 bps/Hz, and this remains effective even when Bob and Eve are in a close distance of $R = 8$ m. Therefore, both spatial parameter D_1 and resolution B play pivotal roles in security performance, necessitating proper design based on practical scenarios. Building on the results of the diminishing returns observed beyond a certain resolution, a 3-bit setting appears as a practical choice, yielding near-optimal

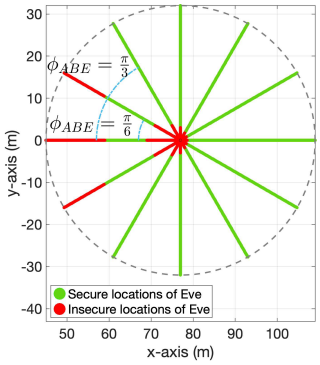


Fig. 7 The secure locations of Eve under different angles.

results compared to an infinite-resolution phase shifter setting.

Furthermore, we investigate the impact of the spatial parameters, including the eavesdropping distance between Eve and Bob R and the receiving distance D_1, on security performance. As delineated in Fig. 6, the secrecy rate increased with the extension of R or a decrease in D_1 attributable to the reduced correlation between legitimate and eavesdropping channels. This observation aligns with Fig. 4, establishing that enhanced spatial correlation reduces secrecy rates. Such an observed trend provides opportunities for maintaining secure transmissions even when distances D_1 exceed 77 m. For systematic analysis on secure locations of Eve, we set a fixed receiving distance $D_1 = 77$ m, and vary the location of Eve around a circle centered at Bob, utilizing predefined angles $\phi_{ABE} \in \{0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{5\pi}{6}, \pi, -\frac{5\pi}{6}, -\frac{2\pi}{3}, -\frac{\pi}{2}, -\frac{\pi}{3}, -\frac{\pi}{6}\}$. This configuration facilitates the systematic evaluation of the impact of eavesdropping distance (corresponding to the radius R of the circle) on the secrecy rate. With a threshold secrecy rate defined at 1.0 bps/Hz, we can obtain the secure locations of Eve. As illustrated in Fig. 7, the secure locations of Eve expand as ϕ_{ABE} increases from 0 to π, the angle relative to the line between Bob and Alice. Notably, when Eve is located between Bob and Alice ($\phi_{ABE} = 0$), a minimal secure location range of Eve is observed, confined between

(59 m, 0 m) and (69 m, 0 m). Two factors influence this behavior. First, as Eve moves away from Bob, the spatial correlation of the channels reduces, enhancing the secrecy rate. Then, as Eve moves closer to Alice while away from Bob, its eavesdropping capability increases, eventually offsetting the benefits garnered from spatial channel variations and reducing the secrecy rate below the threshold. In addition, the augmentation of ϕ_{ABE} can diminish the eavesdropping capability, thereby expanding the secure location range of Eve. For instance, when $\phi_{ABE} = \frac{\pi}{6}$, the secure locations of Eve lie between (59.6 m, 10 m) and (70.9 m, 3.5 m), a range nearly 13 m wider than the range at $\phi_{ABE} = 0°$. For angles in the domain $\frac{\pi}{2} \leq \phi_{ABE} \leq \pi$ and $-\frac{\pi}{2} \geq \phi_{ABE} \geq -\pi$, the secure eavesdropping range broadens, extending radially. Figure 7 also shows that the symmetrical nature of the observed trends for both positive ($0 \leq \phi_{ABE} \leq \pi$) and negative angles ($-\pi \leq \phi_{ABE} \leq 0$). Therefore, the spatial topological relationship between network nodes is essential for optimizing system security performance to construct secure transmission.

In addition, we evaluate the robustness of the LCRFP algorithm by comparing it with other benchmarks with channel mismatch. We assume that the statistical CSI error models of both Bob and Eve follow the CSCG distribution and set the relative amount of CSI uncertainty $\epsilon = \epsilon_G = \epsilon_h = \epsilon_g = \epsilon_{di}, i = \{u, e\}$ and $\epsilon = \{0.1, 0.3, 0.6\}$ to characterize the estimated precision, and the transmit power $P = 5$ dBm. Figure 8 compares the LCRFP algorithm with the non-robust algorithm (LCFP algorithm) in terms of secrecy rate versus the iteration number under the imperfect CSI ($\epsilon = 0.1$) setting. It shows that only the LCRFP algorithm achieves almost sure convergence to a solution after a few iterations, demonstrating the robustness of our proposed algorithm. Then, we depict the achievable SR versus different CSI uncertainty in Fig. 9. The non-robust algorithm under the perfect CSI setting is set as the upper bound, and the random phase scheme under the perfect CSI setting is set as the lower bound. It shows that increasing channel uncertainty ϵ leads to more iteration times to obtain the stationary solution by analyzing the achievable SR versus different CSI uncertainty. Furthermore, it can be seen that there has been an increasing secrecy performance loss with the CSI uncertainty increasing compared with the upper bound. However, the secrecy performance of our proposed algorithm is always over the lower bound. In particular, the proposed algorithm can increase up to 54.6% SR compared with the lower bound when CSI uncertainty is as low as 40%, demonstrating that the LCRFP method can mitigate the loss of the secrecy performance caused by the random system state. In addition, there is only 10.2% secrecy performance loss on the LCRFP algorithm when $\epsilon = 10\%$ (with a secrecy rate of 1.07 bps/Hz) compared with the upper bound. Therefore, we can achieve a close solution by the proposed algorithm with imperfect CSI to that in the case of perfect CSI when the estimation error is in a certain region, and the LCRFP algorithm can improve system secrecy performance compared with the non-robust benchmark schemes.

Finally, we delve into a more practical scenario characterized by imperfect CSI and finite-resolution phase shifters, offering a comprehensive understanding of how these factors impact secrecy performance. As presented in Fig. 10, we examine the impact of CSI uncertainty ($\epsilon \in \{0.1, 0.3, 0.6\}$) across various finite-resolution settings B. Compared with the infinite-resolution setting, we observe that secrecy performance loss amplifies as resolution B diminishes. Specifically, a marked decline is evident as resolution transitions from 1-bit to 3-bit, followed by a more

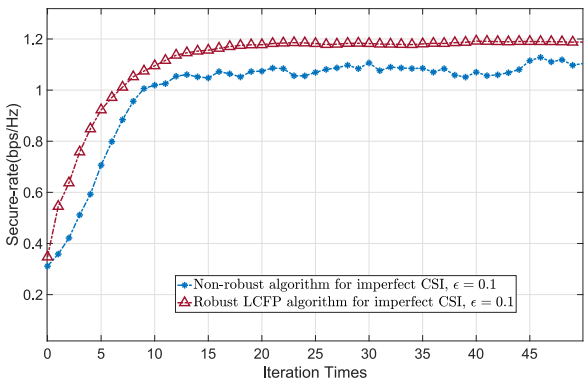


Fig. 8 Convergence of algorithms under imperfect CSI.

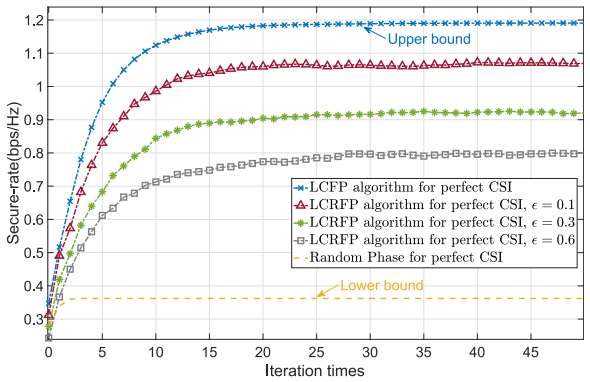


Fig. 9 The LCRPF algorithm versus different CSI uncertainty.

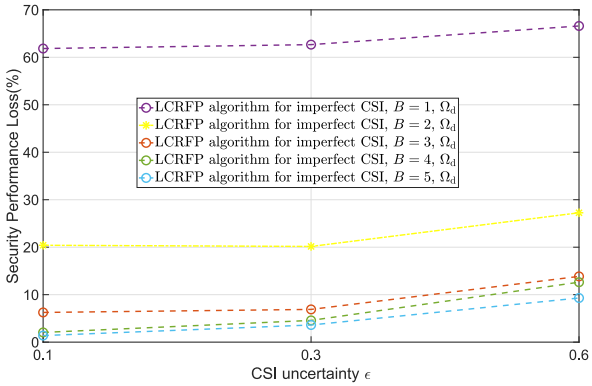


Fig. 10 The performance loss varies with CSI uncertainty under different finite resolution phase shifters compared with infinite resolution phase shifters.

modest reduction as the transition advances from 4-bit to 5-bit. Furthermore, as ϵ increases, the gradual increase in security performance loss becomes evident, which signifies that the SR achieved with a finite-resolution setting gradually falls behind the performance level of the infinite-resolution setting. This trend is attributed to the impact of quantization error, which degrades performance as CSI uncertainty intensifies. Building upon these observations, we provide a threshold for quantitative analysis. Notably, to ensure a security performance loss below 10%, when the channel uncertainty is 0.3 or lower, a choice of $B = 3$ is adequate. However, when the uncertainty escalates to 0.6, a value of $B = 5$ is suggested. Therefore, after making a trade-off, we can obtain that the greater resolution B is necessary to keep the performance loss in a certain range as CSI uncertainty increases. This underscores the practicality of adopting a finite-resolution phase shifter setting in the RIS-aided mmWave system for optimal performance.

7. Conclusion

In this work, we investigated the secure transmission under the effect of the non-neglectable practical issues in a mult-RIS-aided mmWave system, including the high computational complexity, imperfect CSI acquisition, and finite resolution phase shifters limitation. A robust PLS enhancement method was proposed to solve the challenging optimization problem due to the coupled variables, complex functions, non-convex constraints, and uncertain system states. Both theoretical derivations and simulations demonstrated that the proposed algorithms could mitigate the joint negative effects of practical issues and outperform the benchmark schemes in convergence. Moreover, it was illustrated that there exists a tradeoff between secure performance and complexity/overhead, which can help multiple RIS deployments potentially be extended to more practical secure communication scenarios.

Acknowledgments

This work was supported in part by the National Key R&D Program of China (No. 2020YFB1805000), the National Natural Science Foundation of China (NSFC) under Grant 61871070, 91938202, and the Sichuan Science and Technology Program under Grant 2022YFG0005.

References

[1] K. Sakaguchi, G.K. Tran, H. Shimodaira, S. Nanba, T. Sakurai, K. Takinami, I. Siaud, E.C. Strinati, A. Capone, I. Karls, R. Arefi, and T. Haustein, "Millimeter-wave evolution for 5g cellular networks," IEICE Trans. Commun., vol.E98-B, no.3, pp.388–402, March 2015.

[2] K. Ishibashi, T. Hara, S. Uchimura, T. Iye, Y. Fujii, T. Murakami, and H. Shinbo, "User-centric design of millimeter wave communications for beyond 5G and 6G," IEICE Trans. Commun., vol.E105-B, no.10, pp.1117–1129, Oct. 2022.

[3] X. Ge, R. Zi, X. Xiong, Q. Li, and L. Wang, "Millimeter wave communications with OAM-SM scheme for future mobile networks," IEEE J. Sel. Areas Commun., vol.35, no.9, pp.2163–2177, 2017.

[4] A. Almohamad, A.M. Tahir, A. Al-Kababji, H.M. Furqan, T. Khattab, M.O. Hasna, and H. Arslan, "Smart and secure wireless communications via reflecting intelligent surfaces: A short survey," IEEE Open J. Commun. Soc., vol.1, pp.1442–1456, 2020.

[5] A. Mukherjee, S.A.A. Fakoorian, J. Huang, and A.L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," IEEE Commun. Surveys Tuts., vol.16, no.3, pp.1550–1573, 2014.

[6] L. Dong, Z. Han, A.P. Petropulu, and H.V. Poor, "Improving wireless physical layer security via cooperating relays," IEEE Trans. Signal Process., vol.58, no.3, pp.1875–1888, 2010.

[7] J. Zhu, R. Schober, and V.K. Bhargava, "Linear precoding of data and artificial noise in secure massive MIMO systems," IEEE Trans. Wireless Commun., vol.15, no.3, pp.2245–2261, 2016.

[8] E. Tekin and A. Yener, "The general Gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming," IEEE Trans. Inf. Theory, vol.54, no.6, pp.2735–2751, 2008.

[9] L. Hu, H. Wen, B. Wu, F. Pan, R. Liao, H. Song, J. Tang, and X. Wang, "Cooperative jamming for physical layer security enhancement in Internet of things," IEEE Internet Things J., vol.5, no.1, pp.219–228, 2018.

[10] X. Yu, D. Xu, Y. Sun, D.W.K. Ng, and R. Schober, "Robust and secure wireless communications via intelligent reflecting surfaces," IEEE J. Sel. Areas Commun., vol.38, no.11, pp.2637–2652, 2020.

[11] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I.F. Akyildiz, "A new wireless communication paradigm through software-controlled metasurfaces," IEEE Commun. Mag., vol.56, no.9, pp.162–169, 2018.

[12] M. Zhao, Q. Wu, M. Zhao, and R. Zhang, "Exploiting amplitude control in intelligent reflecting surface aided wireless communication with imperfect CSI," CoRR, vol.abs/2005.07002, 2020.

[13] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," IEEE Global Communications Conference, GLOBECOM 2018, Abu Dhabi, United Arab Emirates, pp.1–6, 2018.

[14] P. Wang, J. Fang, X. Yuan, Z. Chen, and H. Li, "Intelligent reflecting surface-assisted millimeter wave communications: Joint active and passive precoding design," IEEE Trans. Veh. Technol., vol.69, no.12, pp.14960–14973, 2020.

[15] H. Shen, W. Xu, S. Gong, Z. He, and C. Zhao, "Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications," IEEE Commun. Lett., vol.23, no.9, pp.1488–1492, 2019.

[16] X. Yu, D. Xu, and R. Schober, "Enabling secure wireless communications via intelligent reflecting surfaces," 2019 IEEE Global Communications Conference, GLOBECOM 2019, Waikoloa, HI, USA, pp.1–6, 2019.

[17] K. Feng, X. Li, Y. Han, S. Jin, and Y. Chen, "Physical layer security enhancement exploiting intelligent reflecting surface," IEEE Commun. Lett., vol.25, no.3, pp.734–738, 2021.

[18] F. Karim, B. Hazarika, S.K. Singh, and K. Singh, "A performance analysis for multi-ris-assisted full duplex wireless communication system," IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2022, Virtual and Singapore, pp.5313–5317, 2022.

[19] W. Mei and R. Zhang, "Performance analysis and user association optimization for wireless network aided by multiple intelligent reflecting surfaces," IEEE Trans. Commun., vol.69, no.9, pp.6296–6312, 2021.

[20] Y. Ogawa, T. Utsuno, T. Nishimura, T. Ohgane, and T. Sato, "Subterahertz MIMO spatial multiplexing in indoor propagation environments," IEICE Trans. Commun., vol.E105-B, no.10, pp.1130–1138, Oct. 2022.

[21] A. Alkhateeb, J. Mo, N. González Prelcic, and R.W. Heath, Jr., "Mimo precoding and combining solutions for millimeter-wave systems," IEEE Commun. Mag., vol.52, no.12, pp.122–131, 2014.

[22] J.A. Zhang, X. Huang, V. Dyadyuk, and Y.J. Guo, "Massive hybrid antenna array for millimeter-wave cellular communications," IEEE Wirel. Commun., vol.22, no.1, pp.79–87, 2015.
[23] S.A. Busari, K.M.S. Huq, S. Mumtaz, L. Dai, and J. Rodriguez, "Millimeter-wave massive MIMO communication for future wireless systems: A survey," IEEE Commun. Surveys Tuts., vol.20, no.2, pp.836–869, 2018.
[24] Y. Cao, T. Lv, and W. Ni, "Intelligent reflecting surface aided multi-user mmWave communications for coverage enhancement," 31st IEEE Annual International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2020, London, United Kingdom, pp.1–6, 2020.
[25] S. Gong, C. Xing, P. Yue, L. Zhao, and T.Q.S. Quek, "Hybrid analog and digital beamforming for RIS-assisted mmWave communications," IEEE Trans. Wireless Commun., vol.22, no.3, pp.1537–1554, 2023.
[26] Y. Xiu, J. Zhao, C. Yuen, Z. Zhang, and G. Gui, "Secure beamforming for multiple intelligent reflecting surfaces aided mmWave systems," IEEE Commun. Lett., vol.25, no.2, pp.417–421, 2021.
[27] K. Yamaguchi, H.P. Bui, Y. Ogawa, T. Nishimura, and T. Ohgane, "Channel prediction techniques for a multi-user MIMO system in time-varying environments," IEICE Trans. Commun., vol.E97-B, no.12, pp.2747–2755, Dec. 2014.
[28] L. Wei, C. Huang, G.C. Alexandropoulos, C. Yuen, Z. Zhang, and M. Debbah, "Channel estimation for RIS-empowered multi-user MISO wireless communications," IEEE Trans. Commun., vol.69, no.6, pp.4144–4157, 2021.
[29] H. Liu, X. Yuan, and Y.A. Zhang, "Matrix-calibration-based cascaded channel estimation for reconfigurable intelligent surface assisted multiuser MIMO," IEEE J. Sel. Areas Commun., vol.38, no.11, pp.2621–2636, 2020.
[30] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "A framework of robust transmission design for IRS-aided MISO communications with imperfect cascaded channels," IEEE Trans. Signal Process., vol.68, pp.5092–5106, 2020.
[31] D. Yang, J. Xu, W. Xu, Y. Huang, and Z. Lu, "Secure communication for spatially correlated RIS-aided multiuser massive MIMO systems: Analysis and optimization," IEEE Commun. Lett., vol.27, no.3, pp.797–801, 2023.
[32] H. Niu, X. Lei, Y. Xiao, M. Xiao, and S. Mumtaz, "On the efficient design of RIS-assisted secure MISO transmission," IEEE Wireless Commun. Lett., vol.11, no.8, pp.1664–1668, 2022.
[33] Z. Zhang, C. Zhang, C. Jiang, F. Jia, J. Ge, and F. Gong, "Improving physical layer security for reconfigurable intelligent surface aided NOMA 6G networks," IEEE Trans. Veh. Technol., vol.70, no.5, pp.4451–4463, 2021.
[34] X. Lu, W. Yang, X. Guan, Q. Wu, and Y. Cai, "Robust and secure beamforming for intelligent reflecting surface aided mmWave MISO systems," IEEE Wireless Commun. Lett., vol.9, no.12, pp.2068–2072, 2020.
[35] L. Zhang, C. Pan, Y. Wang, H. Ren, and K. Wang, "Robust beamforming design for intelligent reflecting surface aided cognitive radio systems with imperfect cascaded CSI," IEEE Trans. Cogn. Commun. Netw., vol.8, no.1, pp.186–201, 2022.
[36] W. Wang, W. Ni, H. Tian, Z. Yang, C. Huang, and K. Wong, "Safeguarding NOMA networks via reconfigurable dual-functional surface under imperfect CSI," IEEE J. Sel. Topics Signal Process., vol.16, no.5, pp.950–966, 2022.
[37] H. Jia, L. Ma, and S. Valaee, "STAR-RIS enabled downlink secure NOMA network under imperfect CSI of eavesdroppers," IEEE Commun. Lett., vol.27, no.3, pp.802–806, 2023.
[38] Z. Li, S. Wang, M. Wen, and Y. Wu, "RIS-aided secure energy-efficiency maximization under uncertain CSI," IEEE Global Communications Conference, GLOBECOM 2022, Rio de Janeiro, Brazil, pp.4637–4642, 2022.
[39] S. Li, B. Duo, M.D. Renzo, M. Tao, and X. Yuan, "Robust secure UAV communications with the aid of reconfigurable intelligent surfaces," IEEE Trans. Wireless Commun., vol.20, no.10, pp.6402–6417, 2021.
[40] C. Liu, C. Tian, and P. Liu, "RIS-assisted secure transmission exploiting statistical CSI of eavesdropper," IEEE Global Communications Conference, GLOBECOM 2021, Madrid, Spain, pp.1–6, 2021.
[41] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," IEEE Trans. Commun., vol.69, no.5, pp.3313–3351, 2021.
[42] G. Zhou, C. Pan, H. Ren, K. Wang, and Z. Peng, "Secure wireless communication in RIS-aided MISO system with hardware impairments," IEEE Wireless Commun. Lett., vol.10, no.6, pp.1309–1313, 2021.
[43] H. Guo, Y. Liang, J. Chen, and E.G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," IEEE Trans. Wireless Commun., vol.19, no.5, pp.3064–3076, 2020.
[44] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," IEEE J. Sel. Topics Signal Process., vol.10, no.3, pp.501–513, 2016.
[45] Q. Wu and R. Zhang, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," IEEE Trans. Commun., vol.68, no.3, pp.1838–1851, 2020.
[46] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H.V. Poor, "Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: Achievable rates with limited discrete phase shifts," IEEE J. Sel. Areas Commun., vol.38, no.8, pp.1809–1822, 2020.
[47] Y. Cao, T. Lv, and W. Ni, "Intelligent reflecting surface aided multi-user mmWave communications for coverage enhancement," 31st IEEE Annual International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2020, London, United Kingdom, pp.1–6, 2020.
[48] H. Zhang, B. Di, L. Song, and Z. Han, "Reconfigurable intelligent surfaces assisted communications with limited phase shifts: How many phase shifts are enough?," IEEE Trans. Veh. Technol., vol.69, no.4, pp.4498–4502, 2020.
[49] A. Almohamad, A.M. Tahir, A. Al-Kababji, H.M. Furqan, T. Khattab, M.O. Hasna, and H. Arslan, "Smart and secure wireless communications via reflecting intelligent surfaces: A short survey," IEEE Open J. Commun. Soc., vol.1, pp.1442–1456, 2020.
[50] Z. Li, W. Chen, Q. Wu, H. Cao, K. Wang, and J. Li, "Robust beamforming design and time allocation for IRS-assisted wireless powered communication networks," IEEE Trans. Commun., vol.70, no.4, pp.2838–2852, 2022.
[51] M.R. Akdeniz, Y. Liu, M.K. Samimi, S. Sun, S. Rangan, T.S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," IEEE J. Sel. Areas Commun., vol.32, no.6, pp.1164–1179, 2014.
[52] D. Shiu, G.J. Foschini, M.J. Gans, and J.M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," IEEE Trans. Commun., vol.48, no.3, pp.502–513, 2000.
[53] K. Xu, J. Zhang, X. Yang, S. Ma, and G. Yang, "On the sum-rate of ris-assisted MIMO multiple-access channels over spatially correlated rician fading," IEEE Trans. Commun., vol.69, no.12, pp.8228–8241, 2021.
[54] H. Gao, P.J. Smith, and M.V. Clark, "Theoretical reliability of MMSE linear diversity combining in Rayleigh-fading additive interference channels," IEEE Trans. Commun., vol.46, no.5, pp.666–672, 1998.
[55] A.D. Dabbagh and D.J. Love, "Multiple antenna MMSE based downlink precoding with quantized feedback or channel mismatch," IEEE Trans. Commun., vol.56, no.11, pp.1859–1868, 2008.
[56] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: channel estimation and reflection optimization," IEEE Wireless Commun. Lett., vol.9, no.4, pp.518–522, 2020.
[57] S.P. Boyd and L. Vandenberghe, Convex Optimization, Cambridge University Press, 2014.

[58] K. Shen and W. Yu, "Fractional programming for communication systems — Part II: Uplink scheduling via matching," IEEE Trans. Signal Process., vol.66, no.10, pp.2631–2644, 2018.
[59] D.P. a. Bertsekas, Nonlinear Programming, 2nd ed., Athena Scientific Belmont, Athena Scientific, MA, USA, 1999.
[60] Y. Xu and W. Yin, "A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion," SIAM J. Imaging Sci., vol.6, no.3, pp.1758–1789, 2013.
[61] M. Razaviyayn, M. Hong, and Z. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," SIAM J. Optim., vol.23, no.2, pp.1126–1153, 2013.
[62] A. Liu, V.K.N. Lau, and M. Zhao, "Online successive convex approximation for two-stage stochastic nonconvex optimization," IEEE Trans. Signal Process., vol.66, no.22, pp.5941–5955, 2018.
[63] W. Shi, X. Zhou, L. Jia, Y. Wu, F. Shu, and J. Wang, "Enhanced secure wireless information and power transfer via intelligent reflecting surface," IEEE Commun. Lett., vol.25, no.4, pp.1084–1088, 2021.

Qingqing Tu received the B.Eng. and M.Eng. degrees in Computer Science and Engineering from the University of Electronic Science and Technology of China (UESTC) in 2011 and 2014, respectively. She has also completed an international exchange study in the University of Electro-Communications in Japan. She is currently working toward the Ph.D. degree in cyberspace security with the University of Electronic Science and Technology of China and serving with the National Computer Network Emergency Response Technical Team/Coordination Center of China. Her research interests include physical layer security and millimeter wave communications.

Zheng Dong received the B.Sc. (in Electronic Information Science and Engineering) and M.Eng. (in Communication and Information System) degrees from the School of Information Science and Engineering, Shandong University, Jinan, China, in 2009 and 2012, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering (in Telecommunications), McMaster University, Canada, in 2016. He was a Postdoc Research Fellow in the School of Electrical and Information Engineering, The University of Sydney, Australia. He is currently a Professor in the School of Information Science and Engineering, Shandong University, China. His research interests include the Industrial Internet of Things and Ultra-reliable Low-Latency Communications.

Xianbing Zou received the B.Eng. degree in Physics at Gannan Normal University, Ganzhou, China, in 1995, and M.Eng. degree in Electromagnetic Field and Microwave Technology at the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 1998. He is now an associate professor of Electronics and Communication Engineering at the National Key Laboratory of Science and Technology on Communications, UESTC. His current research interests are in RF front-end and hardware platform of wireless communication system.

Ning Wei received the B.Eng. and Ph.D. degrees in electrical engineering from the University of Electronic Science and Technology of China (UESTC) in 2003 and 2008, respectively. He was a Visitor of The University of Texas at Austin, Austin, TX, USA, in 2012. He is currently a Full Professor with the University of Electronic Science and Technology of China. His research interests include wireless communications networks and signal processing, including ad hoc networks, multiRAT, channel coding, millimeter wave communications, cooperative communications, and physical-layer network coding. He has served as a Reviewer for various international journals and conferences, including the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

IEICE TRANS. COMMUN., VOL.E107–B, NO.6 JUNE 2024

PAPER

Federated Deep Reinforcement Learning for Multimedia Task Offloading and Resource Allocation in MEC Networks*

Rongqi ZHANG[†,††], ***Nonmember***, **Chunyun PAN**[†,††a)], ***Member***, **Yafei WANG**[†,††], **Yuanyuan YAO**[†,††], ***and*** **Xuehua LI**[†,††], ***Nonmembers***

SUMMARY With maturation of 5G technology in recent years, multimedia services such as live video streaming and online games on the Internet have flourished. These multimedia services frequently require low latency, which pose a significant challenge to compute the high latency requirements multimedia tasks. Mobile edge computing (MEC), is considered a key technology solution to address the above challenges. It offloads computation-intensive tasks to edge servers by sinking mobile nodes, which reduces task execution latency and relieves computing pressure on multimedia devices. In order to use MEC paradigm reasonably and efficiently, resource allocation has become a new challenge. In this paper, we focus on the multimedia tasks which need to be uploaded and processed in the network. We set the optimization problem with the goal of minimizing the latency and energy consumption required to perform tasks in multimedia devices. To solve the complex and non-convex problem, we formulate the optimization problem as a distributed deep reinforcement learning (DRL) problem and propose a federated Dueling deep Q-network (DDQN) based multimedia task offloading and resource allocation algorithm (FDRL-DDQN). In the algorithm, DRL is trained on the local device, while federated learning (FL) is responsible for aggregating and updating the parameters from the trained local models. Further, in order to solve the not identically and independently distributed (non-IID) data problem of multimedia devices, we develop a method for selecting participating federated devices. The simulation results show that the FDRL-DDQN algorithm can reduce the total cost by 31.3% compared to the DQN algorithm when the task data is 1000 kbit, and the maximum reduction can be 35.3% compared to the traditional baseline algorithm.

key words: *multimedia transmission, computing offloading, resource allocation, federated learning, deep reinforcement learning*

1. Introduction

In recent years, with the continuous development of 5G networks, the number of multimedia services and smart terminals in mobile networks has increased rapidly, leading to a significant increase in mobile data volume [1]. According to Cisco's latest forecast report [2], much of the new significant traffic will originate from mobile multimedia services, which will rapidly increase as a percentage of total traffic due to the sheer volume of data. In 2017, mobile multimedia services accounted for 59% of all mobile data traffic. By 2023, this figure will jump to 79%.

Ultra-low latency, intensive computing and massive transmission are the distinctive features of most mobile multimedia services, for example, webcasting, virtual reality services (VR), augmented reality services (AR), cloud computers, and online games. These mobile multimedia services often have high requirements on network latency, bandwidth and computing power. Meanwhile, due to the amount of traffic and computing on mobile users and devices increasing dramatically, multimedia devices need to handle many intensive mobile multimedia tasks such as video compression and transcoding [3], [4]. The huge amount of computation caused by intensive computing tasks puts a lot of pressure on users. However, due to the limited computing resources and storage capacity of multimedia devices, these devices cannot handle tasks locally with low latency as well as low power consumption.

In view of these problems, Mobile Cloud Computing (MCC) has been proposed as a solution, where large amounts of data are centralized in cloud servers to alleviate the burden on local devices [5]. However, traditional cloud computing suffers from problems such as high latency, high load, and core network congestion, as cloud servers are typically deployed at a distance from multimedia devices. In contrast, Mobile Edge Computing (MEC) offers a promising approach by deploying edge servers at edge nodes or base stations. This enables mobile terminals to offload their computing tasks to nearby edge nodes for processing, using wireless channels to reduce task processing delays, improve network utilization efficiency, and enhance Quality of Service [6]. Nevertheless, compared to cloud computing, edge computing is limited by offload decisions, wireless resources and computing resources. Wireless resources mainly include bandwidth and transmitting power. Computational resources generally refer to the CPU frequency of local mobile devices and edge servers. To fully leverage the advantages of the MEC paradigm, there is a need for joint optimization of offloading decisions, communication, and computational resources. This presents a major problem in wireless networks between user devices and MEC servers.

To address this challenge, several studies have inves-

Manuscript received July 5, 2023.
Manuscript revised October 23, 2023.
Manuscript publicized January 30, 2024.

[†]The authors are with the Key Laboratory of Information and Communication Systems, Ministry of Information Industry, Beijing Information Science and Technology University, Beijing 100101, China.

[††]The authors are with the Key Laboratory of Modern Measurement and Control Technology, Ministry of Education, Beijing Information Science and Technology University, Beijing 100101, China.

*This work is partially supported by Beijing Natural Science FoundationHaidian Original Innovation Joint Fund (No. L212026), R&D Program of Beijing Municipal Education Commission (KM202211232011), Key Project of Beijing Natural Science Foundation-Haidian Original Innovation Joint Fund (No.L222004).

a) E-mail: chunyupan@bistu.edu.cn

DOI: 10.23919/transcom.2023EBP3116

tigated the joint allocation of wireless and computational resources in MEC systems [7]–[11]. The Lyapunov optimization methods, online dynamic task scheduling, and game theory has been proposed to solve the problems of joint wireless resources, computational resources, and offloading decisions. The authors in Ref. [7] proposed a local compressed offload model to solve the resource allocation problem of multi-user mobile edge computing offload systems. In Ref. [8], the authors proposed a Lyapunov optimization based approach to study the task assignment scheduling scheme for maximum power consumption and execution delay in MEC systems with energy harvesting capability. The authors in Ref. [9] considered a heuristic algorithm for solving joint resource allocation decisions to minimize the time delay. The authors in Ref. [10] investigated a computational resource allocation scheme based on potential game theory to reduce the energy consumption of MEC networks and improve the efficiency of computational resources. In Ref. [11], the paper proposed a suboptimal resource allocation algorithm that generates priorities for users based on their channel gain and locally calculated energy consumption, and implements different offloading schemes for different priorities to minimize the weighted sum of delay and energy consumption. However, these algorithms are usually time-consuming and computationally intensive in complex MEC networks because they need to constantly resolve the problem in a time-varying MEC network environment.

Deep reinforcement learning has become a trend as an approach in solving optimization problems in MEC systems, in recent years [12]–[17]. DRL can adjust its strategy in unstable environments and can adapt to complex MEC scenarios by making different actions with its intelligences. The DRL agents can make adaptive offloading decisions and resource allocation through the different actions it makes. In Ref. [12], the authors proposed a distributed machine learning approach that makes it possible for DRL to perform online offloading in an MEC environment. The authors in Ref. [13] considered a DRL-based video offload scheme to maximize its long-term performance. The authors in Ref. [14] studied a temporal attentional deterministic policy gradient based on a deep reinforcement learning algorithm called Deep Deterministic Policy Gradient (DDPG) to solve the joint optimization problem of computational offloading and resource allocation in MEC. Ref. [15], this paper proposed a DRL-based offloading scheme to enhance the utility of multimedia devices in dynamic MEC. Simulation results demonstrate that the DRL scheme reduces energy consumption, computational experiments and task failure rate. The authors in Ref. [16] proposed a DRL-based offloading framework that can be adaptive to the common patterns behind various applications to infer the optimal offloading strategy for different scenarios. Ref. [17], the authors propose an advanced deep learning based computational offloading algorithm for multistage vehicle edge cloud computing networks to minimize the total time and energy cost of the whole system. Although DRL is very resilient in complex MEC networks, because most DRL learn in a centralized manner, the required action space and configuration of parameters explode when multimedia devices are added, which directly leads to less efficient training and easier privacy disclosure. To solve this problem, Federated learning (FL) is proposed to optimize MEC networks [18].

Federated learning is a distributed machine learning that enables distributed multiple device nodes to co-communicate and participate in the aggregation of global models. Different devices can perform local model training separately, communicate with each other through federated learning and upload model parameters from local model training for global model aggregation. Federated learning allows the exchange of model parameters without sharing raw data and enhances the collaboration capability of multiple distributed devices and protects the privacy and security of the devices.

Several studies have investigated the resource allocation and computational offloading problems involved in FL for two optimization objectives based on system latency and energy consumption minimization [19], [20]. The authors in Ref. [19] minimized the value of the FL loss function by optimizing the joint resource allocation and UE selection, and satisfied both the latency and energy consumption requirements for performing FL. The authors in Ref. [20] proposed an alternative directional algorithm formulating the joint optimization of CPU frequency and power control as a nonlinear programming (NLP) problem to solve the problem of minimizing the energy consumption of all multimedia devices subject to federated learning time requirements. References [21]–[23] focus on the combined learning of federation learning and deep reinforcement learning, i.e., training local DRL models and then integrating them together to develop a comprehensive global DRL model. The authors in Ref. [21] proposed a joint optimization scheme for optimal path selection and power allocation based on the federal deep Q-network learning algorithm, which maximizes network throughput while ensuring power constraints and mobility constraints, taking into account communication resources, but without considering a reasonable allocation of computational resources. In [22], this paper considered a multimodal deep reinforcement learning framework based on hybrid policies and proposes an online joint collaboration algorithm in combination with FL and validates the performance of the algorithm, however, the intelligent body agent in this work does not undertake some resource allocation operations such as allocation of power, computational offloading of tasks. The authors in [23] proposed a federate cooperative caching framework based on deep reinforcement learning but the work did not take into account task offloading.

We compare the objectives and resource optimization of ourstudy with some related work in the MEC systems, the resultsof which are shown in Table 1. It is obvious that our study can overcome the shortcomings of many previous works.

For mobile multimedia devices, their limited computing resources and battery capacity may hinder efficient task

Table 1 Comparing with some related work.

Reference	Delay	Energy	Communication	Computation	Offloading	DRL	FL
[7]	✓	×	✓	×	✓	×	×
[8]	✓	✓	✓	×	✓	×	×
[9]	✓	×	✓	✓	✓	×	×
[10]	×	×	×	✓	×	×	×
[11]	✓	✓	×	✓	✓	×	×
[12]	×	×	×	✓	✓	✓	×
[13]	✓	×	×	×	✓	✓	×
[14]	✓	✓	✓	×	✓	✓	×
[15]	×	×	×	×	✓	✓	×
[16]	✓	×	✓	×	✓	✓	×
[17]	✓	✓	✓	×	✓	✓	×
[19]	×	×	✓	×	×	×	✓
[20]	✓	✓	×	✓	×	×	✓
[21]	×	×	×	✓	×	✓	✓
[22]	✓	×	✓	×	×	✓	✓
[23]	✓	×	✓	×	×	✓	✓
Our Work	✓	✓	✓	✓	✓	✓	✓

completion. In such cases, offloading tasks to edge or cloud servers becomes necessary. The offloading decision made by the multimedia device plays a critical role in controlling the overall MEC system overhead and ensuring a good user experience. Additionally, task offloading consumes wireless channel resources, necessitating reasonable allocation of these resources in MEC systems. In this paper, we propose an adaptive offloading framework based on federated deep reinforcement learning to jointly optimize transmit power, computational resources, and offloading decisions, with the aim of minimizing delay and energy consumption for mobile multimedia devices in task completion. The contributions of this paper can be summarized as follows:

1. We transform the optimization problem into a multi-objective optimization problem with the objective of minimizing the weighted sum of delay and energy consumption required by the system to perform the task. To solve this complex problem, we jointly allocate computational and communication resources and transform the nonlinear planning problem into a federated deep reinforcement learning problem for multiple intelligent agents.

2. For multimedia devices, the changing location and different kinds of multimedia task of devices cause non-IID data. To reduce the impact of non-IID data. In this paper, we propose a mechanism for selection of participating federal learning devices. To ensure the communication overhead as well as convergence of FL learning.

3. We design an adaptive offloading algorithm based on FL and DRL, which jointly allocates computational resources and task offloading, which not only increases the overall scalability of the system but also accelerates the learning speed of deep reinforcement learning. It maintains relatively stable performance in the complex MEC network environment and outperforms other DRL algorithms.

The rest of this article is organized as follows: Section 2 describe the system model and the problem formulation is described in Sect. 2. In Sect. 4, we present the design of the FDRL-DDQN algorithm. Section 5 presents simulation results. Finally, Sect. 6 shows the conclusion of this paper.

2. System Model

In this paper, we consider a MEC network configuration that consists of a MEC server, an MCC server, a MEC base station (BS) and a set of $\mathcal{N} = \{1, 2, \ldots, N\}$ multimedia devices. As depicted in Fig. 1, when a task is generated, the user's task request is initially submitted to a multimedia device. Subsequently, the MEC BS receives offloading tasks from the multimedia devices. The computing tasks offloaded to the BS are processed by the BS server, and the results are then returned to the terminal. Meanwhile, the remaining multimedia tasks are executed locally. We make a diagram to explain the function of each layer in Fig. 2, and some key parameters are listed in Table 2.

We consider the time into consecutive time frames, which are divided into $\mathcal{T}$ time slots denoted as a set of $\mathcal{T} = \{1, 2, \ldots, T\}$. This article explores a ternary offloading strategy. Specifically, the local offloading decision of device i as x_i, where $x_i = 1$ signifies that the multimedia device executes the tasks locally, and $x_i = 0$ means the multimedia device offloads the multimedia tasks to the MEC server or MCC server. Moreover, we use y_i and z_i to represent the multimedia devices' offloading of computing tasks to the MCC server and MEC server, respectively. In this context, $z_i = 1$ denotes offloading to the MEC server, while $y_i = 1$ implies offloading to the MCC server. Therefore, we have the ternary offloading strategy as follows:

$$x_i + y_i + z_i = 1, \ \forall i \in \mathrm{N}. \tag{1}$$

2.1 Computing Model

This section focuses on modeling the delay and energy con-

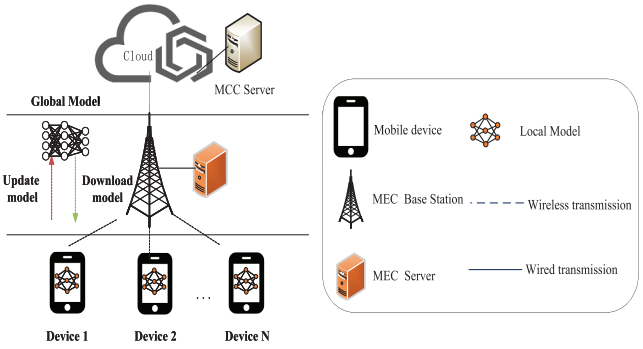


Fig. 1 A MEC system model.

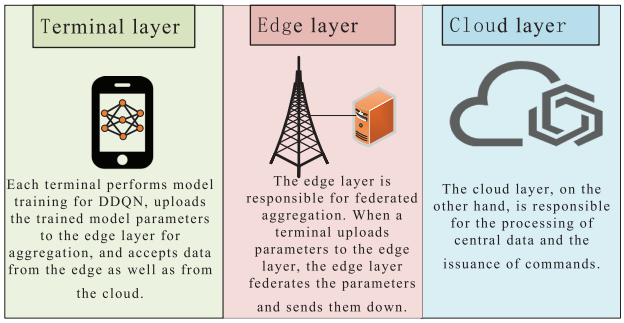


Fig. 2 The function of each layer.

Table 2 Key parameters.

Notation	Meaning
L_i	The size of tasks
x_i, y_i, z_i	binary offloading variables
r_i	the uplink transmission rate of multimedia device i
${T_i}^{bs}$	The communication delay of mobile task offloading
${E_i}^{bs}$	The energy consumption of mobile task offloading
${T_i}^{e}$	The computation delay of MEC server
${T_i}^{c}$	The computation delay of MCC server
${T_i}^{E}$	The delay of tasks offloading to the MEC server
${T_i}^{C}$	The delay of tasks offloading to the MCC server
${T_i}^{L}$	The processing dealy for task on device
${E_i}^{L}$	The processing energy consumption for task on device
θ^{eval}	The eval network parameter
θ^{global}	The global model parameter
$Q_i(s,a)$	The action state values of Bellman equation
$L(\theta)$	Loss function

sumption experienced by multimedia devices during the execution of multimedia tasks. When multimedia device i offloads its task to either the MEC server or MCC server, various factors come into play, including the size of the multimedia tasks, channel conditions, and transmitting power. It's important to note that the transmission of offloaded multimedia tasks occurs over wireless channels, involving both multimedia devices and base stations. Furthermore, the execution of multimedia tasks requires the allocation of uplink frequency resources for transmission. Therefore, the uplink transmission rate of multimedia device i is determined by

$$r_i = B\log_2(1 + \frac{p_i h_i}{\sigma^2}), \tag{2}$$

where B and p_i denote the operating bandwidth and transmit power of the multimedia device, respectively, h_i and σ^2 denote the transmission link gain and channel noise between the multimedia device and the base station, respectively.

The communication delay and the energy consumption of mobile task offloading are respectively given by

$${T_i}^{bs} = \frac{L_i}{r_i}, \tag{3}$$

$${E_i}^{bs} = p_i {T_i}^{bs}, \tag{4}$$

where L_i is the size of task (in bit). Since the computing power of edge servers and clouds is very resource-rich compared to local devices, this paper ignores the computing consumption at the edge servers or clouds, so only the energy consumed by their task transmission is calculated. From the device's point of view, when a task is offloaded to either server, the energy consumed to process the task is the energy spent on the task transfer. So both of the energy consumption of MCC server and MEC server would be equal to E_i^{bs}.

Due to limited computing power and battery capacity, multimedia devices offload tasks to edge servers or cloud services to meet QoS requirements. The computation delay of MEC server and MCC, while offloading is given, respectively, as follows:

$$T_i^e = \frac{C_i}{F^e}, \tag{5}$$

$$T_i^c = \frac{C_i}{F^c}, \tag{6}$$

where F^e and F^c denote the average computing power of the edge server and the cloud, respectively. C_i denotes the CPU cycle requirement of the task (in cycle/second). The delay of the multimedia tasks offloading to the MEC server and MCC server respectively as follows:

$$T_i^E = T_i^e + T_i^{bs}, \tag{7}$$

$$T_i^C = T_i^c + T_i^{bs}. \tag{8}$$

Assuming that user-submitted multimedia tasks are selected for execution on the local multimedia device and they do not need to be offloaded to the edge server for processing, the processing delay and energy consumption for task on device is defined as

$$T_i^L = \frac{C_i}{f_i^L}, \tag{9}$$

$$E_i^L = \kappa_i (f_i^L)^2, \tag{10}$$

where κ_i is the energy consumption factor related to the multimedia device, which depends on the CPU performance architecture of the terminal.

In this paper, we aim to optimize the computational resource allocation as well as the offloading policy, which minimizes the multimedia task execution cost. The long-term expected cost of each multimedia device is a weighted

sum of execution delay and energy consumption. Each of multimedia device cost is given by

$$T_i(p_i, f_i, x_i, y_i, z_i) = x_i T_i^L + y_i T_i^E + z_i T_i^C, \tag{11}$$

$$E_i(p_i, f_i, x_i, y_i, z_i) = x_i E_i^L + y_i E_i^e + z_i E_i^c, \tag{12}$$

where p_i, f_i, x_i, y_i and z_i represent the vectors of transmit powers, computation resource allocation, local computing, edge offloading, and cloud offloading decision of device i, respectively.

3. Problem Formulation

In this paper, the problem is formulated to solve the joint minimization of long-term delay and energy consumption of multimedia devices over time $\mathcal{T}$.

In solving the optimization problem of offloading decisions and computational resource allocation for MEC systems, the objective of this paper is to minimize the total cost of the combination of execution delay and energy consumption of the devices in the MEC system. Based on the above analysis, the optimization problem can be described as follows:

$$\begin{aligned}
&\min_{p_i, f_i, x_i, y_i, z_i} \quad \omega T_i + \lambda E_i \\
&\text{subject to :} \\
&C1 : f_i^L \le F_{\max}, \forall i \in N \\
&C2 : x_i E_i^L + y_i E_i^e + z_i E_i^c \le E_{\max,i}, \forall i \in N \\
&C3 : T_i \le T_{\max}, \forall i \in N \\
&C4 : x_i + y_i + z_i = 1, \forall i \in N \\
&C5 : x_i, y_i, z_i \in \{0, 1\}, \forall i \in N,
\end{aligned} \tag{13}$$

where ω and λ in the above optimization problem are denoted as the delay and energy consumption weighting factors of device i in performing the multimedia task, respectively. Let $0 \le \omega \le 1$, and $0 \le \lambda \le 1$, $\omega + \lambda = 1$, the ratio of ω to λ is a constant, the value of the weight factor should be chosen according to the heterogeneity of the resources available on each multimedia device, if the device receives greater constraints in terms of energy resources than computational resources, the value should be larger, otherwise it should be smaller. The constraint C1 indicates that the computing resources allocated for the user should not exceed the total computing capacity of the MEC system F_{max}. C2 indicates a limit on the energy resources of the device, which should not exceed the maximum energy E_{max} that the MEC system can provide, and C3 expresses that the overall service time cost should not exceed the maximum allowable delay for the user T_{max}. C4 and C5 are the ternary offloading schemes used in this paper.

However, to satisfy the requirement of minimizing the total system cost under the multimedia task execution delay as well as energy consumption tolerance. With binary offloading variables (x_i, y_i, z_i) included of above formulated problem (13) makes the problem into a mixed integer nonlinear programming (MINLP) problem that cannot be solved in an acceptable time frame.

4. The Proposed FDRL-DDQN Algortihm

In this section, we present our solution to address the complex and non-convex optimization problem. We propose a deep reinforcement learning algorithm that combines federated learning, and for offloading actions, we adopt the Dueling DQN algorithm. This algorithm is referred to as FDRL-DDQN.

The framework of the FDRL-DDQN algorithm is illustrated in Fig. 3. The FDRL-DDQN algorithm contains three main components: the training of offloading decision and resource allocation, federated aggregation and update of local model parameters. In the first step, devices participating in federated learning are selected. Next, the local model is trained to learn multimedia task offloading decisions and resource allocation. Subsequently, the trained model parameters are federated and aggregated. Finally, the updated parameters are distributed to each multimedia device involved in federated learning. Algorithm 1 provides a detailed description of the proposed FDRL-DDQN algorithm in this paper. And We give a flow chart of the Federation framework in Fig. 4.

In a complex MEC network environment, mobile multimedia devices are faced with three options for computing multimedia tasks. This results in a total of 2^{3N} possible computation offloading options per device at each time slot. With an increasing number of multimedia devices, the complexity of the state and action spaces for intelligent agents also grows exponentially. Consequently, implementing centralized training becomes extremely challenging when dealing with large-scale datasets and expansive action spaces. Moreover, in mobile multimedia services, which involve extensive data transmission, centralized training leads to significant

Algorithm 1 The FDRL-DDQN algorithm

Input: wireless channel gains h_i, size of multimedia tasks L_i.
Output: offloading action α_i, resource allocation action f_i.
Set the total time frame t, maximum FLDDQN iteratons to M.
Initialize the networks parameters $\theta_i^{eval} = \theta_i^{target}$ of all device and global model parameters θ^{global}.
while $(M >= 0)$: **do**
  Set $\theta_i^{eval} = \theta_i^{target} = \theta^{global}$
  Select the set of participating training devices, $\mathcal{W}$,
  for $t = 1, 2, ..., T$ **do**
    Offloading action $\alpha_i = f_i(h_i, L_i)$
    Compute Q_i for all α_i
    Select the optimal offloading action $\hat{\alpha}_i = arg\ max Q_i$
    for each device i in I **do**
      Interact with environment and calculate the cost
      Train the model and get θ_i^{eval}
      Upload the network parameters θ_i^{eval} of the terminal device i to the MEC server
      Federated average use by (20) in MEC server and get the θ^{global}.
      Transmits θ^{global} to the devices to replace the original network parameters θ_i^{eval}.
    end for
  end for
end while

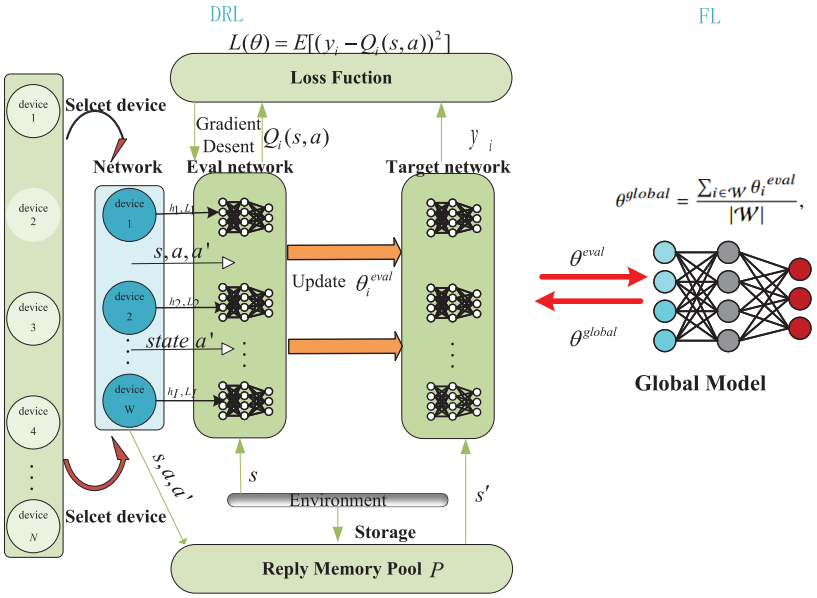


Fig. 3 Federated framework model.

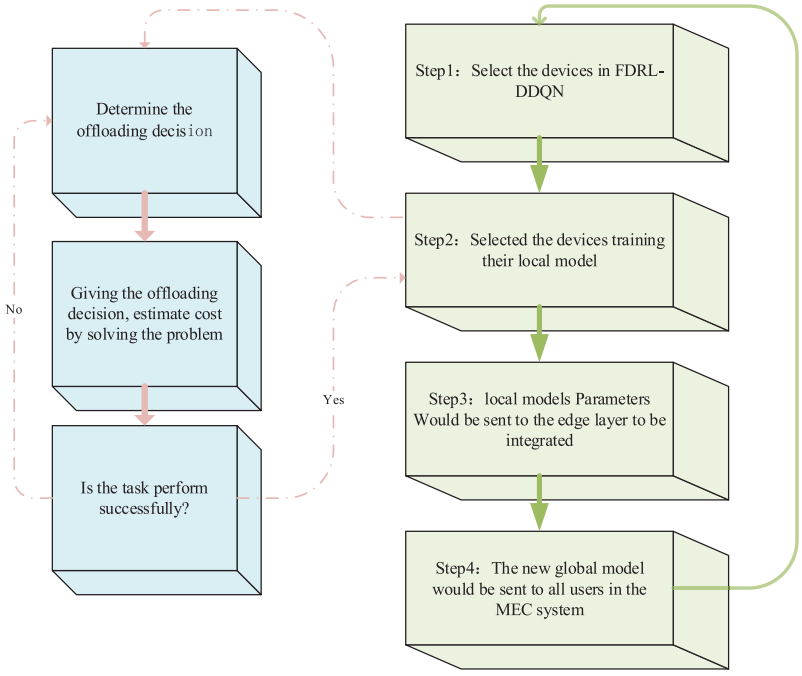


Fig. 4 Federated framework model.

communication overhead and raises privacy concerns. Federated learning, a distributed machine learning approach, offers several advantages in the MEC environment. Firstly, it enables individual training of intelligent agents, allowing them to cooperate and make independent decisions during multimedia task execution. This approach enhances learning efficiency, reduces communication overhead, and better adapts to large-scale MEC networks. Secondly, federated learning facilitates interactive updates of model parameters between distributed and central nodes, eliminating the need for sharing original data. This mechanism provides a robust guarantee for the security of local data.

4.1 Select Device

In cases where a large number of devices are involved in joint learning, it can result in increased drop rates and unnecessary communication overhead. To address this issue,

we introduce a device selection strategy in this paper. At the beginning of each iteration of the FDRL-DDQN algorithm, a specific set of multimedia device agents is carefully chosen to participate in the learning process. The principles of how to select devices in this article are as follows

$$\arg\max_{i\in N} \mathrm{MSE}(\frac{d_i P_{\max,i}}{F_{\max,i}}), \tag{14}$$

where d_i denotes the distance between the device and the BS and the function MSE denotes the mean squared error. This approach allows for the selection of the most cost effective method for different multimedia devices. Choosing the right device to participate in the learning process can also help with the overall learning speed.

4.2 Local Model Training

For the training of local agents in the FDRL-DDQN algorithm, each multimedia device employs the DDQN algorithm to train its own local model and learn offloading and resource allocation strategies. The Dueling DQN algorithm utilizes an experience pool to store data for each state at time t. This includes the action chosen based on the current state, the reward received for that action, the new state after performing the action, and whether the state terminates during the training process. As the amount of stored data in the experience pool reaches a sufficient size, small batches of data are randomly selected and fed into the neural network for training. This continuous training process optimizes the weight parameters in the neural network. By randomly selecting training data, a broad range of experiences can be learned, breaking the correlation between sample data and preventing overfitting issues resulting from local experiences. Based on the system model and optimization objectives presented in this chapter, the FDRL-DDQN algorithm defines three key elements: the state space, action space, and reward function, which can be defined as

1. State space:

   $$S_i = \{L_i, h_i\}, \tag{15}$$

   where S_i denotes the state space of each device, L_i denotes the amount of multimedia task, and h_i denotes the path gain of infinite transmission between the multimedia device and the base station.

2. Action space:

   In the FDRL-DDQN model considered in this paper, the intelligence is responsible for making appropriate decisions based on the computational multimedia tasks. The decisions include, determining whether the computational multimedia tasks are offloaded to the edge server or the cloud server, and how much computational resources should be allocated when the multimedia tasks are executed locally. The action space consists of two parts, the multimedia device offloading decision $\{\alpha_i^L, \alpha_i^E, \alpha_i^C\}$, where α_i^L denotes that the task is executed locally, α_i^E denotes that the task is offloaded to the edge server, and α_i^C denotes that the task is offloaded to the cloud. The resource allocation strategy $f = \{f_1, f_2, \ldots, f_N\}$.

3. Reword function:
   The cost of each agent is the weighted sum of the delay and energy consumption in the objective function. The optimization objective of this paper is to minimize the cost, so the reward function should be negatively correlated with the cost, so the reward function as shown

   $$R_i = -\left(\frac{\omega(y_i T_i^E + z_i T_i^C) + \lambda(y_i E_i^e + z_i E_i^e)}{\omega x_i T_i^L + \lambda x_i E_i^L}\right). \tag{16}$$

The Dueling DQN algorithm is utilized to address complex decision control challenges in real-world multimedia environments. It combines Q-learning algorithms, empirical replay mechanisms, and target Q-values based on action value functions to approximate the Q-value of the optimal policy. Q-learning selects the action with the highest Q-value by consulting the Q-table, while dueling DQN uses a neural network to obtain the corresponding Q-value based on the input, resulting in improved operational speed and stability. As depicted in Fig. 5, the Dueling DQN architecture divides the fully connected layer of the network into two branches, each with its specific output. The upper branch represents the state value function, which quantifies the value of the static state environment itself, irrespective of actions taken. The lower branch represents the state-dependent action advantage function, which captures the average action payoff relative to states, indicating the additional value brought by decision-making behavior. These two branches are then combined to derive the Q-value for each action. This approach allows for mutual supervision, eliminates redundant degrees of freedom, mitigates the risk of inflated Q-value estimates, and enhances algorithm stability. Therefore, in this paper, the notation $u_i(s, a)$ is employed to represent the direct cost incurred by each device as determined through the aforementioned optimization process. Using the Bellman equation, the action state values are given by

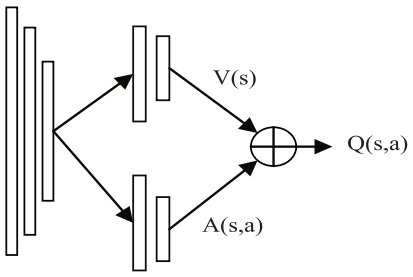


Fig. 5 The network of DDQN algorithm.

$$Q_i(s,a) = u_i(s,a) + \gamma \sum_{s' \in S} P_{ss'}(a) \max Q_i(s',a'), \quad (17)$$

where $Q_i(s,a)$ denotes the Q value corresponding to the action a generated according to the current state s. Similarly, $\max Q_i(s',a')$ is used to denote the action a' corresponding to the maximum Q value output by the stare s'. $P_{ss'}(a)$ stands for the transfer probability function, γ represent the discount factor. There are two network parameters in the dueling DQN network, one is the current Q network parameter, denoted by θ_i^{eval}, to evaluate the greedy strategy of the current network and the other is the target network parameter, denoted by θ_i^{target}, to evaluate the target value y_i. In each training iteration, the target value used to train the evaluation network in device i is calculated as

$$y_i = u_i(s,a) + \gamma Q_i(s', \arg\max_{a' \in A} Q_i(s',a';\theta_i^{eval}), \theta_i^{t\arg et}). \quad (18)$$

Meanwhile, to obtain the optimal strategy and minimize the gap between the target value and the evaluated value, we set the loss function as

$$L(\theta) = E[(y_i - Q_i(s,a))^2]. \quad (19)$$

4.3 Parameter Aggregation and Update

At the start of each learning round in FDRL-DDQN, the participating local devices upload their network parameter models to the MEC server for model aggregation. This aggregation process combines the models to create a global model within the MEC. Subsequently, the MEC server distributes the aggregated global model parameters to each multimedia device participating in FDRL-DDQN as the network parameters for the next round. In this paper, we employ FedAvg [24] as the model aggregation method.

$$\theta^{global} = \frac{\sum_{i \in \mathcal{W}} \theta_i^{eval}}{|\mathcal{W}|}, \quad (20)$$

where $|\mathcal{W}|$ denotes the total number of participating training devices. θ^{global} represents the global model parameters. Once the global model aggregation is complete, the model parameters are transmitted to the local device, which then updates its own model parameters. The local device utilizes its local data to train the network evaluation parameters during the offloading decision update. This update process continues iteratively for the local devices until the algorithm converges.

5. Simulation Results

In this section, we use tensorflow1.0 GPU version to implement the FDRL-DDQN framework in python and perform simulations to evaluate its performance. The main simulation parameter settings in this paper are shown in Table 3.

To simulate the proposed FDRL-DDQN algorithm, we constructe a network comprising 50 multimedia devices. However, only 10 devices were selected for each training round. Each device has a maximum computational capacity of 1 Gbps and a maximum energy consumption of 23 dBm. Due to the limitations of computing power on the multimedia devices in the MEC network, we utilize the smallest feasible neural network for our algorithm. Considering the computational constraints, our neural network consisted of an input layer, two hidden layers, and an output layer. The first and second layers consisted of 32 and 16 neurons, respectively. ReLU activation functions were used throughout the network.

Table 3 Simulation parameters.

Parameter	Value
Moblie device transmit power (p_i)	23 dBm
The number of CPU cycles required to process 1bit data (c)	500
Channel noise power (σ^2)	10^{-9}
Computing power of edge server and cloud server(F^e), (F^c)	6, 7, 8, 9, and 10 GHz
Computing power of the multimedia device (f_i^L)	500 KHz
Communication bandwidth (B)	180 KHz
Greedy coefficient (ε)	0.95
Discount factor (γ)	0.9
Learning rate	0.001
$T_{\max}$	0.5 s
Experience replay buffer	3000
Channel gain (h_i)	$[1,2] * 10^{-5}$ Uniform distribution

5.1 Convergence Performance

In this section, we evaluate the convergence performance of the FDRL-DDQN algorithm and compare it with the distributed DDQN. We examine the convergence of the two schemes using selected devices and all devices to participate in the federation to address the non-IID data issue of multimedia devices. Additionally, we analyze the impact of learning rate and batch size on the convergence of the FDRL-DDQN algorithm.

Firstly, we assess the convergence speed of the training loss in the FDRL-DDQN algorithm. In Fig. 6, the average training loss $L(\theta)$ of the FDRL-DDQN model is plotted. Initially, the algorithm exhibits significant fluctuations due to the lack of experience during the initial training, making it challenging for the intelligence to learn the optimum. However, as the experience pool accumulates sufficient data, the intelligence can make actions that lead to the optimal solution, resulting in maximum rewards. As the number of iterations increases, the loss function steadily decreases, indicating a smoother learning process. After approximately 200 iterations, the algorithm reaches the optimum value for the neural network.

In Fig. 7, we address the non-IID problem among mobile multimedia devices by selecting only a fraction of devices to participate in each round of federated learning, ensuring the convergence speed of the overall algorithm. To

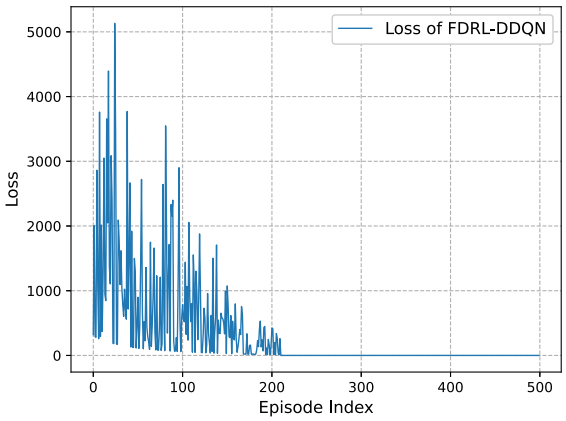


Fig. 6 Convergence process of loss function.

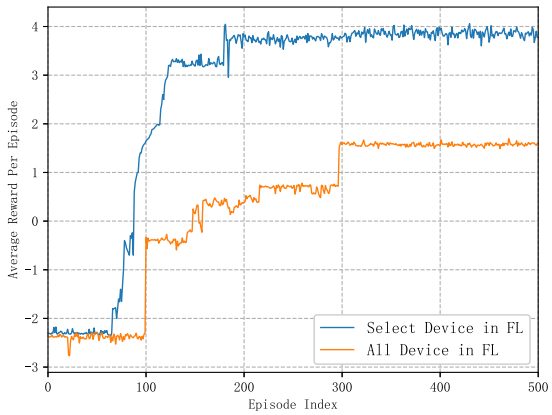


Fig. 7 Comparison of convergence between FDRL-DDQN and systems without selection mechanism.

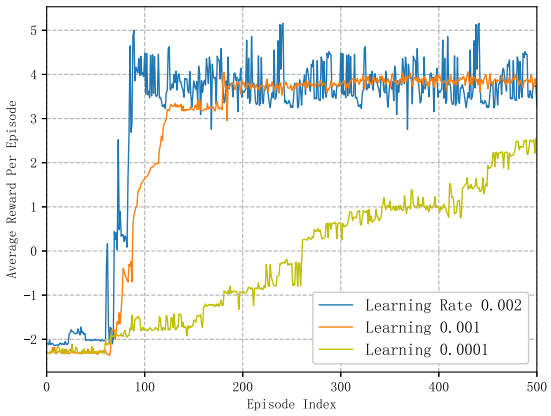


Fig. 8 Convergence of reward value under different learning rates.

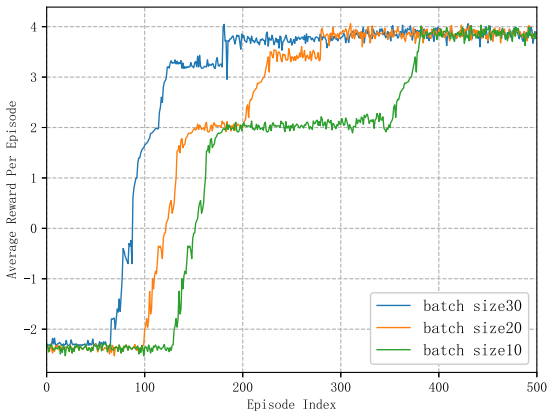


Fig. 9 Convergence of reward value under different batch size.

evaluate the effect of the device selection method on the convergence of the FDRL-DDQN algorithm, we compare the approach of adding all devices to federated learning with the partial device addition. The convergence performance of the FDRL-DDQN algorithm is validated using 5W randomly generated multimedia tasks, which are offloaded for optimal resource allocation decisions. In each iteration round, the intelligence derives a reward value based on its decision. The average reward of the FDRL-DDQN algorithm is plotted, showing an increasing trend with the number of iterations as the intelligence improves its decision-making ability. The algorithm converges after approximately 200 iterations. It can be observed from the figure that the average reward value of the overall algorithm, after utilizing the device selection mechanism, is significantly higher and converges faster compared to the approach of adding all devices to federated learning.

Furthermore, we examine the impact of learning rate and batch size on the convergence of the FDRL-DDQN algorithm. Figure 8 illustrates the effect of the intelligence's learning rate on the convergence performance of the FDRL-DDQN framework. We experiment with learning rates of 0.0001, 0.001, and 0.002. While a higher learning rate leads to faster learning, the figure shows that a learning rate of 0.0001 results in slow convergence due to the low learning rate. On the other hand, a learning rate of 0.002 increases learning efficiency, but it compromises algorithm stability, causing repeated oscillations that hinder convergence. Thus, in this paper, we set the learning rate to 0.001 in the simulation.

Another parameter of interest is the batch size for multimedia task processing. Figure 9 demonstrates that increasing the batch size improves the convergence of the FDRL-DDQN algorithm. With a small batch size of 10, convergence takes around 380 iterations. However, when the batch size is increased to 20, convergence occurs after 310 iterations, and with a batch size of 30, the algorithm converges quickly in only 180 iterations. Larger batch sizes enable training with more instances, providing the intelligence with faster experience accumulation. Consequently, the executed actions can reach optimal solutions more rapidly, resulting in faster algorithm convergence.

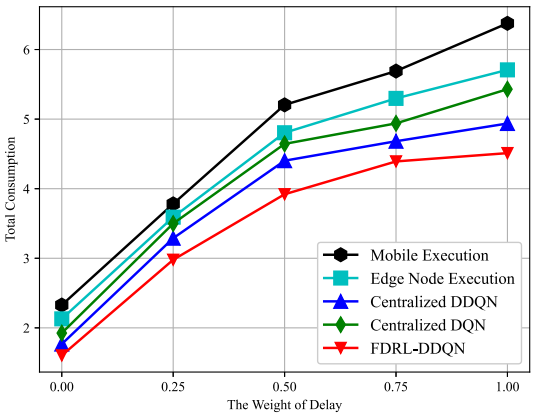


Fig. 10 Influence of delay weight on total system cost.

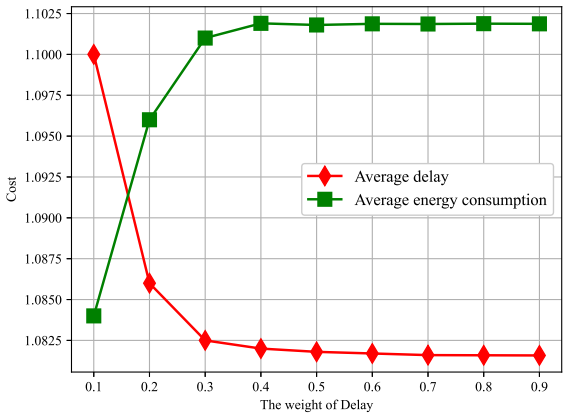


Fig. 11 The average delay and the average energy consumption curve.

5.2 Comparison of Total Cost

In this section, we first compare the proposed FDRL-DDQN algorithm with the distributed DDQN algorithm. To further evaluate the algorithm's performance, we also compare it with the centralized DDQN algorithm, centralized DQN algorithm, and two baseline computation offloading policies: mobile execution and edge node execution. Mobile execution refers to local device computation of multimedia tasks, while edge node execution involves offloading all tasks to the edge node. We investigate the effect of delay weights on the FDRL-DDQN algorithm in comparison to the four mentioned centralized algorithms. Additionally, we discusse the trade-off between delay and energy consumption.

In Fig. 10, we experiment with the weights of delay and energy consumption on algorithm performance. To handle different types of mobile multimedia tasks, we set the weights of delay ω to equal values of 0-1 and similarly γ to $1-\omega$. We set the delay weights of the centralized DQN algorithm, DDQN algorithm, to be consistent with FDRL-DDQN. As ω increases the system assembly also rises, the total cost of FDRL-DDQN is always lower than the other four algorithms. This is due to the fact that FDRL-DDQN is able to provide an optimal offloading strategy for the optimized target compared to the other four algorithms, thus achieving an overall cost reduction.

Figure 11 illustrates the equilibrium trend of the network's average delay and average energy consumption as the delay weight varies. In this simulation, the number of users is set to N=5. From Fig. 11, it is apparent that the network's average delay gradually decreases as the delay weight increases, while the average energy consumption of the network increases. Both the average delay and average energy consumption of the network stabilize when the delay weight reaches a certain threshold when the $\omega >= 0.4$. This happens because when the value of ω is small, increasing it reduces delay at the expense of energy performance. However, when the value of ω is large, due to limitations on user transmitting power, further increasing ω doesn't result in reduced average delay or increased average energy consumption.

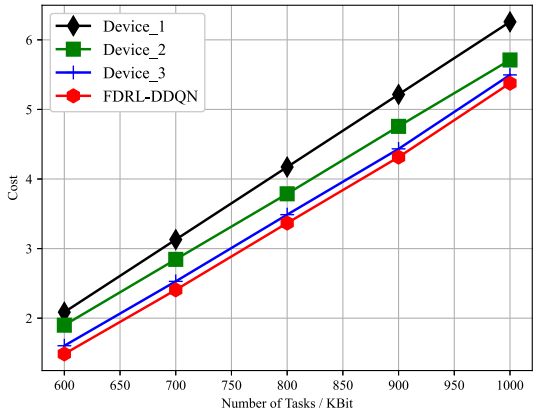


Fig. 12 The cost of FDRL-DDQN scheme is compared with distributed DDQN algorithm.

Figure 12 compares the average total cost of the proposed FDRL-DDQN algorithm with the distributed DDQN algorithm. We select three multimedia devices which are trained individually using the distributed DDQN algorithm without any parameter exchange between the three multimedia devices during the period. When the training is finished, we add the three multimedia devices to the FDRL-DDQN framework for retraining until convergence. The experimental results show that the cost of each device is reduced using the FDRL-DDQN algorithm, where the average consumption can be reduced by 20.3%. By combining federated learning with deep reinforcement learning, cooperative training between devices is achieved. It avoids the equipment alone training by environmental instability, action space, state space and other inexperienced impact, and provides a relatively stable intelligent body learning environment, combining different devices together intelligently. Because of the devices involved in training only upload the model parameters needed for learning, it can effectively protect the privacy and security of users. In addition, federated learning enables knowledge sharing between devices and

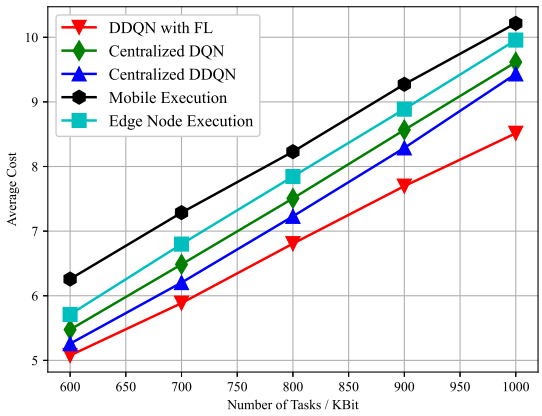


Fig. 13 Effect of the number of multimedia tasks on the total cost of the system.

enriches the data parameters they can collect.

To demonstrate the performance benefits of the FDRL-DDQN algorithm, in Fig. 13, it shows the impact of multimedia tasks data size on the average total cost of devices. We compared FDRL-DDQN algorithm with centralized DDQN algorithm, centralized DQN algorithm, mobile execution algorithm and edge note execution algorithm. It can be seen that the FDRL-DDQN has a faster learning speed than the basic centralized DDQN, and the cost rises more slowly with the number of tasks, with the smallest optimization performance at a total cost of 600 Kbits and the largest optimization performance at a total cost of 1000 Kbits. This is due to the fact that when the multimedia task data size increases, more and more information be interacted between devices, the advantage of federated learning can be fully reflected, and the learning speed of the intelligence will be faster and faster, while the centralized DDQN will decrease due to the increase of the multimedia task data size, which leads to the exponential growth of the action space received by the intelligence. FDRL-DDQN algorithm can reduce the total cost by 7.1% when the system multimedia task volume is 600 Kbit and 31.3% when the multimedia task volume is 1000 Kbit compared with centralized DDQN, while FDRL-DDQN algorithm can reduce up to 35.3% compared with mobile local offloading algorithm and 34.8% compared with edge node offloading algorithm, because Federated Deep Reinforcement Learning combines FL and DRL are organically combined to obtain the exact optimal policy by intelligent and effective learning for multiple device parameters. In addition, FDRL-DDQN can reduce the cost by up to 30.1% compared to the centralized DQN algorithm which has already widely used to offload multimedia task policies. The biggest advantage of the DDQN algorithm over the DQN algorithm is that it can ensure the stability of the target network, which helps the whole system to update the parameters and thus converge faster to get the optimal policy. When FL is combined with DDQN its effect is more obvious, FL makes DDQN algorithm more stable making the final result more accurate and faster convergence.

6. Conclusion

In this paper, we propose an adaptive offloading algorithm FDRL-DDQN that combines federation learning and deep reinforcement learning. For the computational offloading problem in the MEC scenario of mobile multimedia dynamic multimedia task arrival, we jointly allocate computational and communication resources with the goal of minimizing latency and energy consumption, and make reasonable offloading decisions. For the non-IID data problem with different multimedia devices, we design an adaptive device selection mechanism as a way to ensure the convergence of FL. In addition, we compared FDRL-DDQN with centralized Dueling DQN, distributed DDQN, mobile algorithm and edge algorithm with better results. Simulation results show that the algorithm has good latency and energy performance. In future work, we will consider resource coordination among multiple MEC servers, as well as investigate more flexible and generalized resource allocation and computational offloading strategies.

References

[1] A. Nauman, Y.A. Qadri, M. Amjad, Y.B. Zikria, M.K. Afzal, and S.W. Kim, "Multimedia internet of things: A comprehensive survey," IEEE Access, vol.8, pp.8202–8250, 2020.
[2] C.V.N. Index, "Global mobile data traffic forecast update," Cisco White Paper [Online]. Available: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.pdf, 2014.
[3] D. Xu, Q. Li, and H. Zhu, "Energy-saving computation offloading by joint data compression and resource allocation for mobile-edge computing," IEEE Commun. Lett., vol.23, no.4, pp.704–707, 2019.
[4] J. Ren, G. Yu, Y. Cai, and Y. He, "Latency optimization for resource allocation in mobile-edge computation offloading," IEEE Trans. Wireless Commun., vol.17, no.8, pp.5506–5519, 2018.
[5] M. Muniswamaiah, T. Agerwala, and C.C. Tappert, "A survey on cloudlets, mobile edge, and fog computing," 2021 8th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2021 7th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom), pp.139–142, 2021.
[6] X. Wang, Y. Han, C. Wang, Q. Zhao, X. Chen, and M. Chen, "In-edge AI: Intelligentizing mobile edge computing, caching and communication by federated learning," IEEE Netw., vol.33, no.5, pp.156–165, 2019.
[7] J. Ren, G. Yu, Y. Cai, Y. He, and F. Qu, "Partial offloading for latency minimization in mobile-edge computing," GLOBECOM 2017 - 2017 IEEE Global Communications Conference, pp.1–6, 2017.
[8] G. Zhang, W. Zhang, Y. Cao, D. Li, and L. Wang, "Energy-delay tradeoff for dynamic offloading in mobile-edge computing system with energy harvesting devices," IEEE Trans. Ind. Informat., vol.14, no.10, pp.4642–4655, 2018.
[9] Y. Deng, Z. Chen, and X. Chen, "Resource allocation for multi-user mobile-edge computing systems with delay constraints," GLOBECOM 2020 - 2020 IEEE Global Communications Conference, pp.1–6, 2020.
[10] H. Liu, H. Jia, J. Chen, X. Ge, Y. Li, L. Tian, and J. Shi, "Computing resource allocation of mobile edge computing networks based on potential game theory," 2018 IEEE 4th International Conference on Computer and Communications (ICCC), pp.693–699, 2018.
[11] C. You, K. Huang, H. Chae, and B.H. Kim, "Energy-efficient resource allocation for mobile-edge computation offloading," IEEE Trans.

Wireless Commun., vol.16, no.3, pp.1397–1411, 2017.
[12] L. Huang, S. Bi, and Y.J.A. Zhang, "Deep reinforcement learning for online computation offloading in wireless powered mobile-edge computing networks," IEEE Trans. Mobile Comput., vol.19, no.11, pp.2581–2593, 2020.
[13] T. Zhao, L. He, X. Huang, and F. Li, "DRL-based secure video offloading in MEC-enabled IoT networks," IEEE Internet Things J., vol.9, no.19, pp.18710–18724, 2022.
[14] J. Chen, H. Xing, Z. Xiao, L. Xu, and T. Tao, "A DRL agent for jointly optimizing computation offloading and resource allocation in MEC," IEEE Internet Things J., vol.8, no.24, pp.17508–17524, 2021.
[15] J. Wang, J. Hu, G. Min, W. Zhan, Q. Ni, and N. Georgalas, "Computation offloading in multi-access edge computing using a deep sequential model based on reinforcement learning," IEEE Commun. Mag., vol.57, no.5, pp.64–69, 2019.
[16] B. Guo, X. Zhang, Y. Wang, and H. Yang, "Deep-Q-network-based multimedia multi-service QoS optimization for mobile edge computing systems," IEEE Access, vol.7, pp.160961–160972, 2019.
[17] M. Khayyat, I.A. Elgendy, A. Muthanna, A.S. Alshahrani, S. Alharbi, and A. Koucheryavy, "Advanced deep learning-based computational offloading for multilevel vehicular edge-cloud computing networks," IEEE Access, vol.8, pp.137052–137062, 2020.
[18] J. Konen, H.B. Mcmahan, F.X. Yu, P. Richtárik, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," arXiv preprint, arXiv:1610.05492v1, 2016.
[19] M. Chen, Z. Yang, W. Saad, C. Yin, H.V. Poor, and S. Cui, "A joint learning and communications framework for federated learning over wireless networks," IEEE Trans. Wireless Commun., vol.20, no.1, pp.269–283, 2020.
[20] J. Yao and N. Ansari, "Enhancing federated learning in fog-aided iot by CPU frequency and wireless power control," IEEE Internet Things J., vol.8, no.5, pp.3438–3445, 2020.
[21] R. Luo, H. Tian, and W. Ni, "Communication-aware path design for indoor robots exploiting federated deep reinforcement learning," 2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), pp.1197–1202, IEEE, 2021.
[22] Z. Zhu, S. Wan, P. Fan, and K.B. Letaief, "Federated multiagent actor–critic learning for age sensitive mobile-edge computing," IEEE Internet Things J., vol.9, no.2, pp.1053–1067, 2021.
[23] X. Wang, C. Wang, X. Li, V.C. Leung, and T. Taleb, "Federated deep reinforcement learning for internet of things with decentralized cooperative edge caching," IEEE Internet Things J., vol.7, no.10, pp.9441–9455, 2020.
[24] H.B. McMahan, E. Moore, D. Ramage, and B.A. y Arcas, "Federated learning of deep networks using model averaging," arXiv preprint, arXiv:1602.05629, vol.2, 2016.

Rongqi Zhang is currently pursuing the M.Phill degree with the School of Information and Communication Engineering, Beijing Information Science and Technology University. He research interests include mobile edge computing, reinforcement learning and resource allocation.

Chunyu Pan received the Ph.D. degree with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China. She is currently an Associate Professor in the School of Information and Communication Engineering at Information and Communication Engineering. Her research interests include cell free networks, heterogeneous wireless networks, resource allocation and mobile edge computing.

Yafei Wang received the Ph.D. degree in telecommunications engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 2013. She is currently a Professor in Beijing Information Science and Technology University. His main research interests include signal integrity analysis, and RF communication circuit design.

Yuanyuan Yao received the Ph.D. degree in information and communication engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2017. Since 2017, she has been with the School of Information and Communication Engineering, Beijing Information Science and Technology University, Beijing, China, as an associate professor. Her research interests include Cooperative communication of UAV; RIS assisted communication; Stochastic geometry and its applications in large-scale wireless networks; Intelligent Radio Resource Allocation in 6G etc.

Xuehua Li received the Ph.D. degree in telecommunications engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 2008. She is currently a Professor and the Deputy Dean of the School of Information and Communication Engineering with Beijing Information Science and Technology University, Beijing. She is a Senior Member of the Beijing Internet of Things Institute. Her research interests are in the broad areas of communications and information theory, particularly the Internet of Things, and coding for multimedia communications systems.

IEICE TRANS. COMMUN., VOL.E107–B, NO.6 JUNE 2024

PAPER

LSTM Neural Network Algorithm for Handover Improvement in a Non-Ideal Network Using O-RAN Near-RT RIC

Baud Haryo PRANANTO[†a)], ***Student Member***, **ISKANDAR**[†b)], **HENDRAWAN**[†c)], ***and*** **Adit KURNIAWAN**[†d)], ***Nonmembers***

SUMMARY Handover is an important property of cellular communication that enables the user to move from one cell to another without losing the connection. It is a very crucial process for the quality of the user's experience because it may interrupt data transmission. Therefore, good handover management is very important in the current and future cellular systems. Several techniques have been employed to improve the handover performance, usually to increase the probability of a successful handover. One of the techniques is predictive handover which predicts the target cell using some methods other than the traditional measurement-based algorithm, including using machine learning. Several studies have been conducted in the implementation of predictive handover, most of them by modifying the internal algorithm of existing network elements, such as the base station. We implemented a predictive handover algorithm using an intelligent node outside the existing network elements to minimize the modification of the network and to create modularity in the system. Using a recently standardized Open Radio Access Network (O-RAN) Near Realtime Radio Intelligent Controller (Near-RT RIC), we created a modular application that can improve the handover performance by determining the target cell using machine learning techniques. In our previous research, we modified The Near-RT RIC original software that is using vector autoregression to determine the target cell by predicting the throughput of each neighboring cell. We also modified the method using a Multi-Layer Perceptron (MLP) neural network. In this paper, we redesigned the neural network using Long Short-Term Memory (LSTM) that can better handle time series data. We proved that our proposed LSTM-based machine learning algorithms used in Near-RT RIC can improve the handover performance compared to the traditional measurement-based algorithm.

key words: *cellular, handover, 5G, LTE, machine learning, lstm, neural network*

1. Introduction

Handover is one of the crucial processes in cellular communication especially in high mobility users such as vehicular terminals. This process can be defined as the process that prevents ongoing communication from getting interrupted as the mobile equipment changes its attachment point such as cells [1]. However, some disruptions may occur in active communication due to packet losses and delays and these disruptions may result in significant loss of performance [2]. In the 5G era, this handover process is getting more crucial due to the usage of a higher frequency spectrum [3] that causes a smaller cell range.

The traditional handover algorithm is usually reliable in the ideal network condition. However, in some non-ideal network conditions, such as the presence of a coverage hole, this algorithm may not be reliable and result in transmission failure. We prove this through our simulation described in Sect. 8.

Apart of the traditional handover algorithm that will be described in Sect. 2, many other algorithms are proposed including the machine-learning-based algorithms [4] (described in Sect. 3). Neural networks are one of the most popular machine-learning-based methods for handover improvement [5]–[10]. The main issue with those proposed methods is their real-world implementation because machine learning is not originally part of the cellular networks [11]. Most of the machine learning algorithms to improve handover performance require major modifications in the existing cellular networks for their implementation. This will raise many problems in the network deployment and implementation stage.

O-RAN Alliance consortium [12] introduces the Near Real Time Radio Intelligent Controller (Near-RT RIC), a new additional network element in the radio access network (RAN) that can host applications to control base stations. Using this Near-RT RIC, a machine learning algorithm for improving the handover process can be implemented modularly without major modifications to the existing cellular networks. Our implementation of Near-RT RIC will be further described in Sect. 4.

In our previous papers, we described the implementation of machine learning in Near-RT RIC to take advantage of its modularity aspect. The machine learning algorithm can be implemented modularly outside the base station without modifying the current software of the base station. We proved that this method performs better compared to the traditional handover algorithms in a simulated non-ideal network, in this case, a network with coverage holes. We measured the performance in terms of data transmission (i.e., file download) success rate if the user moves along the network and performed a handover. The target cell is determined by several methods, the traditional algorithm and our proposed machine-learning-based algorithm.

As described in [13], we modified the Near-RT RIC original software to fit our simulation case. We modified the vector autoregression (VAR) algorithm used the original in Near-RT RIC software to consider the UE movement and

Manuscript received August 18, 2023.
Manuscript revised November 19, 2023.
Manuscript publicized January 30, 2024.

[†]The authors are with School of Electrical Engineering and Informatics, Bandung Institute of Technology, Indonesia.
a) E-mail: baud.prananto@students.itb.ac.id
b) E-mail: iskandar@itb.ac.id
c) E-mail: hendrawan@itb.ac.id
d) E-mail: adit@itb.ac.id

DOI: 10.23919/transcom.2023EBP3139

compared the performance of this proposed method with the traditional handover algorithm. In the simulation result, we showed that this method can improve the handover performance in a network with a coverage hole.

In our next publication ([14], we extended our research by replacing the VAR algorithm with a Multi-Layer-Perceptron (MLP) neural network. It is proven that this method is also superior compared to the traditional handover algorithm. However, this simple neural networks still underperformed the VAR method. In this paper, we performed further improvement in the neural network to increase the performance.

1.1 Research Motivation and Contribution

The motivation of this research is to improve the machine learning method implemented in Near-RT RIC to solve the handover reliability issue in a non-ideal network. Using Near-RT RIC, the machine learning algorithm to control the handover process can be implemented modularly without major modification of the existing network elements.

In our previous research, we used a simple MLP neural network and it was still underperformed since it did not consider the time-series nature of the input data. We used the user measurement data as the input to determine the target cell in the handover process. In our MLP design, we statically use several last measurement data as the input and thus cannot consider the temporal feature of the time-series input. We suggest that this was the main cause of the underperformance.

The contribution of this research is designing and implementing a Long Short-Term Memory (LSTM)-based handover algorithm in Near-RT RIC to control the handover process. LSTM is chosen as it is better to handle the time-series data and consider the temporal feature of the data. We consider this time-series data handling for machine learning-based handover as our novel contribution.

In this paper, we performed modifications in Near-RT RIC original software, more precisely in the QoE Predictor xApp. We have done two modifications in our previous research and we will briefly review them in Sect. 5: adapt the VAR (the original algorithm used in Near-RT RIC) to consider the UE movement and replaced the vector autoregression with MLP neural network. Our newly proposed method to improve our previous ones is the LSTM neural network that will be described in Sect. 6. We have done simulations to test the performance of those methods and we also studied the effect of training data amount on the handover performance (described in Sect. 7). Finally, we compared the performance of our proposed methods with the traditional handover algorithm and showed that the machine-learning-based handover in Near-RT RIC performs better in a non-ideal condition, in this case, a network with a coverage hole (described in Sect. 8).

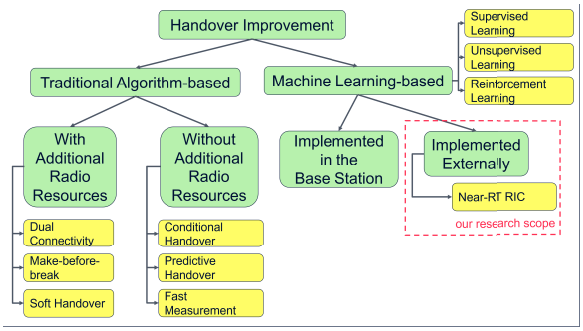


Fig. 1 Solution taxonomy for handover improvement.

1.2 Research Scope and Limitation

This research focuses on the usage of a machine learning algorithm in Near-RT RIC to control the handover process. In this study, we compare the performance of the handover control in Near-RT RIC with the baseline traditional handover algorithm.

There are various solutions to improve the handover performance and we organized the solutions taxonomy in Fig. 1. We focus only on the machine-learning-based solution that is implemented externally for modularity reasons, and we compare the result of our methods with the baseline traditional handover algorithms. The innovations that are implemented on top of the traditional handover algorithm such as soft handover, conditional handover, make-before-break, are not considered and not compared with our proposed machine-learning-based algorithm in Near-RT RIC. The other state-of-the art machine learning algorithms for handover improvement as described in Sect. 3 are also not compared with our method.

1.3 Paper Organization

In this Sect. 1, we provide a gentle overview of our research, problem statement, motivation, contribution, and the scope of this study.

Section 2 describes the Traditional Handover Algorithm, the baseline algorithm for cellular mobility management that we would like to improve using our research. We present the way of working, the issues, limitations, and room of improvement of this traditional algorithm.

We use machine learning techniques to improve the Traditional Handover Algorithm. In Sect. 3, we describe the State of The Art of the usage of Machine Learning for handover improvement. Here we explored the previous works conducted to improve the handover algorithm using various machine learning techniques.

Our solution utilizes an O-RAN new network element called Near-RT RIC, which is described in Sect. 4. Here we also describe how the software is implemented and used for our solution.

Next in Sect. 5, we reviewed our previous methods that we already published in earlier publications [13], [14]. Here

we explain how we modified the original Near-RT RIC software to utilize our proposed machine learning methods: the modified VAR and MLP Neural Network.

In Sect. 6, we introduce our newly proposed method to be used in Near-RT RIC using LSTM. Here we explain the novel algorithm to improve the handover performance.

To test our proposed method, we designed a simulation that will be explained in Sect. 7. We also explain how we collect and utilize the data to prove the effectiveness of our proposed methods.

The result of our simulation with the proposed methods is discussed in Sect. 8. We show the improvement in handover performance compared to the traditional handover algorithm. The overall conclusion and possible future works are written in Sect. 9.

2. Traditional Handover Algorithm

In the traditional handover algorithm [15], the UE sends measurement reports to the serving base station about the condition of serving cell and neighbor cells. The measurement report is about the cell's signal strength (Reference Signal Received Power - RSRP) and/or signal quality (Reference Signal Received Quality - RSRQ). The serving base station will analyze the measurement report to determine the target cell for the handover destination. Usually the target cell is the best-measured neighbor cell (Fig. 2).

The handover process may interrupt the data transmission because it caused temporary disconnection of the UE from the serving cell (thus stopping the data transmission) and connects again to the target cell. The interruption is defined as Mobility Interruption Time (MIT) and 3GPP defines MIT as the shortest time duration supported by the system during which a user terminal cannot exchange user plane packets with any base station during transitions [17]. MIT can be calculated as [2]:

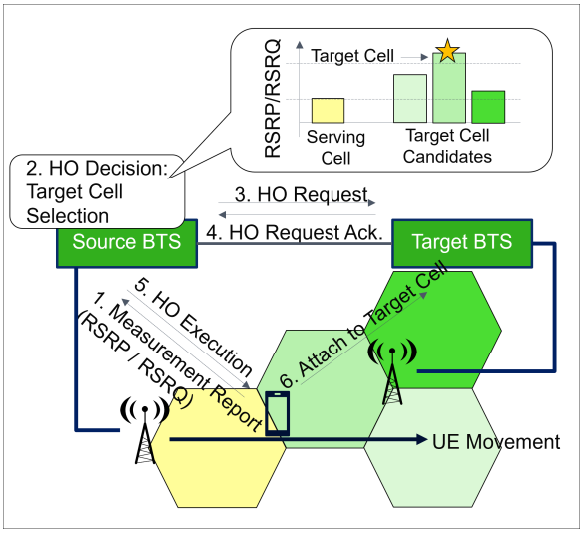


Fig. 2 Traditional handover algorithm [16].

$$T_{\mathrm{MIT}} = \{(1 - P_{\mathrm{HOF}}) \times T_{\mathrm{HIT}}\} + \{P_{\mathrm{HOF}} \times T_{\mathrm{HOF}}\} \quad (1)$$

T_{MIT} = Total MIT
P_{HOF} = Probability of either a handover failure (HOF) or a radio link failure (RLF) during handover
T_{HIT} = Handover Interruption Time, MIT in a successful handover
T_{HOF} = Handover Failure Time, MIT in a HOF or RLF

The T_{HOF} contributes more significantly to MIT (T_{MIT}), thus reducing the T_{MIT} can be better done by reducing P_{HOF}. In LTE Network, T_{HIT} is reported around 50 ms while T_{HOF} ranges from several hundred milliseconds to a few seconds [18]. This means the best way is to avoid unnecessary handovers or handovers to the wrong cell. Target cell determination is very crucial in the handover process to minimize MIT.

The traditional handover algorithm is reliable in ideal conditions, where RSRP/RSRQ measurement always reflects the real condition of the network. Using this algorithm, the best target cell to continue the network connection is always the cell with the best RSRP/RSRQ measurement. In a non-ideal condition, the RSRP/RSRQ measurements may not reflect the real network condition.

An example of this non-ideal network is the presence of a cell coverage hole due to an obstacle. A UE may be handed over to a target cell with the best RSRP/RSRQ, but it enters the target cell's coverage hole after the handover, and the connection fails after that. In this case, the traditional handover algorithm is not reliable to determine the target cell correctly and ensure network connectivity. Our simulation proves the unreliability of the traditional handover algorithm in Sect. 8. This raises the need for machine-learning-based target cell determination.

Besides the baseline traditional handover algorithm, which is part of the base station algorithm as compliance to 3GPP standard [16], there are some other handover algorithm that aims primarily to reduce MIT by reducing P_{HOF}. Some innovations include fast measurements [19], soft-handover, dual connectivity [20], make-before-break [2], [21], conditional handover [5], [22], [23], and predictive handover [5], [24]–[28].

Fast measurement makes the source base station send a handover command before an abrupt deterioration of the radio link to the UE. The UE reacts faster to the channel changes and improves mobility robustness.

Soft handover, dual-connectivity, and make-before-break work similarly by making multiple separate connections to different radio resources simultaneously. This improves mobility robustness but increases the network complexity and requires more radio resources.

In traditional handover, the handover command is sent when the radio conditions start to get degraded [23]. Conditional handover prepares in advance multiple candidate target cells in the network. This enables the handover command to be sent to the UE earlier than at the traditional handover when the radio conditions are still good.

In predictive handover, the candidate target cells are predicted using various techniques, including user behavior and learning the network condition using machine learning techniques.

3. Machine Learning for Handover: State of the Art and Related Works

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed [29]. It studies the computer algorithms that improve automatically through experience [30].

Since the traditional handover algorithm is sometimes not reliable in a non-ideal network condition, some alternative methods are required to determine the target cell, and one of the approaches is the predictive handover using machine learning. Several studies [4] have implemented machine learning to improve handover performance using the predictive handover method (i.e. predict the target cell using machine learning).

Supervised learning is widely used for handover improvement. The neural networks (NN) method is one of the most popular techniques used in several studies [5]–[10]. Some studies use support vector machine [31] and K-nearest neighbor [32], [33]. Unsupervised learning techniques are also used by some studies, for example, K-means [27], [34] and Long Short-Term Memory [25]. Reinforced learning is used by some researchers that usually employ Q-learning algorithms [24], [35].

The neural networks method is popular in mobility management improvement studies. The basic idea behind these studies is to use the concept of neural networks to learn a mobility-based model for every user in the network and then make predictions of which cell the user is most likely to be next [4].

Several previous studies [6], [36] used neural networks for target cell selection in the handover process and performed simulations to justify their proposed method. We based our research on these works and improved them using our proposed methods. The previous research did not perform the implementation of the software in the Near-RT RIC platform and considered only the simple MLP neural network. As a novelty of our research, we present the implementation of the machine learning algorithm in the real Near-RT RIC platform. We also test several methods other than simple MLP neural networks, including the LSTM-based neural network that can better process time-series measurement data.

4. O-RAN Near-RT RIC

Open Radio Access Network (O-RAN) Alliance standardizes and introduces several new applications for open and intelligent RAN on top of the legacy cellular network. This enables the introduction of machine learning applications since machine learning is not originally part of the cellular

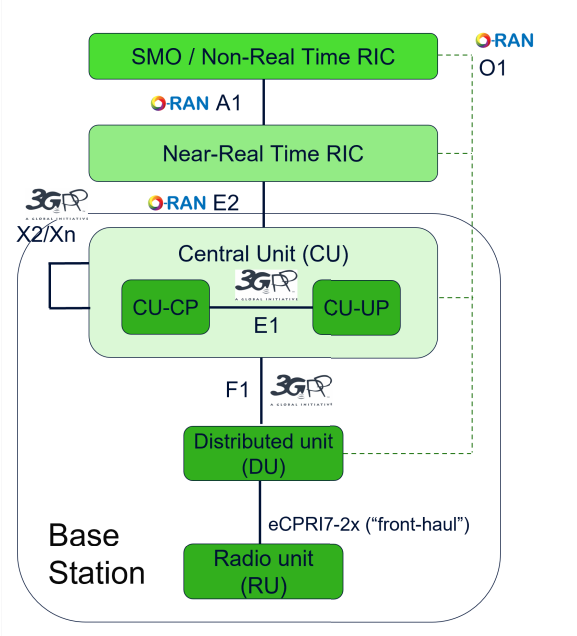


Fig. 3 O-RAN architecture [12].

network standard. O-RAN introduces new network elements called Radio Intelligent Controller (RIC) to add intelligence to the cellular radio network. There are two variants of RIC: Near-Real Time (Near-RT) and Non-Real Time (NRT) RIC (Fig. 3).

Near-RT RIC hosts applications that require real-time response, such as mobility management applications like handover control. Because of this response requirement, Near-RT RIC is typically implemented in an Edge Cloud, a virtual environment that is placed physically near the radio network. The effectiveness of Edge Cloud is already proven to implement RAN elements [37].

NRT-RIC hosts applications that do not require immediate response such as network monitoring and optimization. It can be implemented in Central Cloud and typically colocated with the existing network management system.

There are some use cases defined by O-RAN Alliance [38] to be implemented in O-RAN to provide RAN openness and intelligence, for example, Context-Based Dynamic HO Management for V2X, Flight Path-Based Dynamic UAV Radio Resources Allocation, QoE Optimization, and Traffic Steering. However, the exact implementation of the use case is given to specific vendors. For example, Nokia prioritizes Traffic Steering and Network Anomaly Detection use case for its RIC solution [39].

Several studies already use O-RAN RIC architecture for many applications such as connection management [40], mobility management [41], and scheduling policy optimization [42]. Various machine learning algorithms are implemented in RIC including reinforcement learning [43].

The Near-RT RIC can be implemented in any virtualized environment. In our research, we installed it on an Ubuntu-based virtual machine by installing the open-source

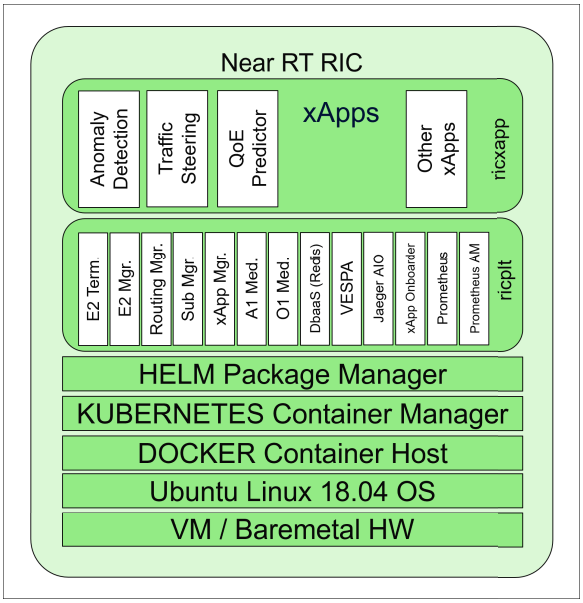


Fig. 4 The software architecture of RIC [44].

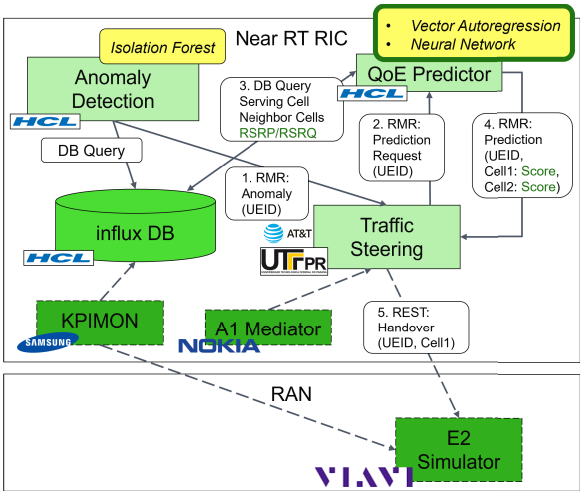


Fig. 5 Anomaly detection use case of Near-RT RIC [45].

software provided by the O-RAN Software Consortium (SC) [44]. The software is container-based and contains several applications called xApps. This architecture can be viewed in Fig. 4.

The Anomaly Detection use case [45] is one of the already existing Near-RT RIC use case examples from O-RAN SC that mostly corresponds to our research need. However, we have to perform some modifications to fit our simulation scenario.

The software contains three xApps: Anomaly Detection, Traffic Steering, and Quality of Experience (QoE) Predictor. The xApps exchange messages in RMR protocol, the Near-RT RIC internal communication. Currently, in this research, the Near-RT RIC works stand alone without any connection to the RAN, and all simulation data is stored in the database.

The scenario begins with the Anomaly Detection xApp detects an anomalous UE, for instance, the UE experiencing degradation of RSRP. In this research, this information is obtained from the database but in the real implementation, this information is notified by RAN (i.e. base station). The Anomaly Detection xApp then informs the anomaly to the Traffic Steering xApp.

Traffic Steering xApp then sends a message to QoE Predictor xApp sending the identity of the UE experiencing an anomaly. QoE Predictor xApp then predicts the score of QoE of the UE if the UE is placed in the neighboring cells. In the original software, this score is the throughput of the data transmission and is predicted using the vector autoregression (VAR) method. Therefore, QoE Predictor predicts the throughput experienced by the UE if it is placed in a certain cell.

This prediction is sent back to Traffic Steering xApp. Based on this prediction, it will perform some necessary actions. The action can be a handover command to the cell where the throughput prediction is the highest one. From this scenario, it is clear that QoE Prediction is the one that actually determines the target cell by performing a prediction of QoE (score) in each cell. The Traffic Steering xApp is just simply choosing the target cell with the highest score.

In this research, we modified the original Near-RT RIC xApps in the Anomaly Detection Use Case to adapt to our simulation scenario. We mainly performed modifications in QoE Predictor xApp as it is the one that actually performs predictions that will determine the target cell. We performed two modifications to the original QoE Predictor xApp. The first modification is to adapt the original software to our simulation scenario. The prediction is still done by the vector autoregression method. The second modification is completely replacing the vector autoregression with a neural network. The neural network design is based on our previous studies [46], [47] that yield optimum results.

5. Our Previous Methods: A Review

The original QoE Predictor xApp software provided by O-RAN SC predicts the QoE using the VAR method. However, the original software is not immediately usable for our research case so we have to perform some modifications in the original xApp. Our research aims to determine the target cell in a non-ideal network containing a coverage hole. This target cell is determined by the movement of the UE that is reflected in the RSRP/RSRQ measurements.

5.1 Modified VAR

Vector Autoregression (VAR) is a statistical time series model used to analyze the relationship between multiple variables. In a VAR model, each variable in the system is modeled as a function of its past values and the past values of all the other variables in the system. A VAR model of order p, denoted as VAR(p), is a set of linear equations that relate each variable in the system to its own past values and

the past values of all the other variables in the system up to p lags. The equations can be written in matrix form as:

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \ldots + A_p Y_{t-p} + u_t \quad (2)$$

where Y_t is a k-dimensional vector of the current values of the k variables in the system, $A_1, A_2, \ldots, A_p$ are $k \times k$ matrices of coefficients that capture the dynamic relationships between the variables at lags 1 to p, and u_t is a k-dimensional vector of error terms that represent the unexplained part of the system at time t.

The original QoE Predictor xApp determines the target cell by predicting the throughput of each cell using time-series throughput data in the training data. However, this software only considers the position of the UE, i.e. what the neighbor cells are. It does not consider the movement and the trajectory angle of the UE. If we use the unmodified original software and training data, the prediction will always give the same target cell for all simulation cases.

Our proposed modified method using VAR can be expressed in the following pseudocode (Algorithm 1). The *italic* expression in Algorithm 1 indicates our modification.

Algorithm 1 Predict the throughput of all cells

Require: list of all cells serving and neighboring UE
and the RSRP measurements of those cells
for all cells in list **do**
Query throughput of cell over time from Training Data
(where RSRP measurement is similar with the one reported by UE)
Remove outliers of the query result
Predict the next throughput of the cell using Vector Autoregression
end for
Report the throughput prediction of all cells in list to Traffic Steering xApp

To adapt the original software to our simulation scenario, we reconstructed the software and training data to put the UE movement into account. The UE movement and its trajectory angle can be reflected by the RSRP measurement variations. From the training data generation process described in Sect. 7, we construct the new training data that considers the UE movement to predict the next throughput by evaluating RSRP values. Using this modified training data, the target cell is determined by the UE movement, not only the UE position like the original QoE Predictor xApp.

5.2 MLP Neural Network

For our second method, we completely replaced the VAR in the QoE Predictor xApp with a neural network. In this preliminary stage, we use a very simple Multi-Layer Perceptron (MLP) neural network regression model to predict whether the download is successful or failed using RSRP and RSRQ samples as input.

This method works in a different approach than the previous one. The VAR method views the problem as a prediction problem, this MLP neural network method views the problem as a classification problem. The prediction problem in the VAR method requires throughput data and determines the target cell by choosing the neighbor cell with the highest predicted throughput. The throughput data is not immediately available in real-life handover cases so we decide to use the already available RSRP/RSRQ data. It is possible to perform prediction of RSRP/RSRQ data but in a non-ideal network, RSRP/RSRQ data do not directly reflects the throughput or QoE of the user.

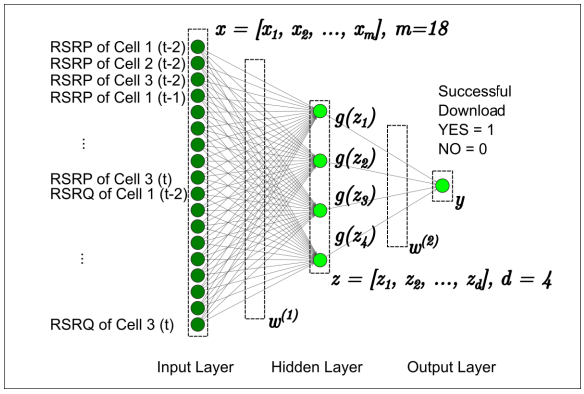


Fig. 6 MLP neural network design.

In our second method, we decided to view the problem as a classification problem, without necessarily predicting the future RSRP/RSRQ data. We directly collect the RSRP/RSRQ data reported by the UE to determine whether it is good or not to perform a handover in a certain neighbor cell.

The MLP neural network in our method contains fully interconnected 18 input nodes, 4 hidden nodes, and 1 output node (Fig. 6). The inputs are the last 3 samples of RSRP and RSRQ measurements from all the 3 cells. The output is whether the data transmission (i.e. file download) is successful or not, represented by the number 0 (failed download) or 1 (successful download). The result of the output node is a floating point continuous number between 0 and 1 that can be used as the prediction score. The score will then be sent to Traffic Steering xApp and the cell with the highest score will be determined as the target cell. In this method, we do not need throughput data and perform any prediction to determine the target cell.

We use Tensor Flow Keras API for the implementation. Currently, the training process is done with 150 times iterations through the whole training data (*epoch = 150*), and the model is updated every 10 training data (*batch size = 10*).

The MLP neural network method has a simpler implementation than the VAR method since it uses only RSRP/RSRQ measurement without a throughput measurement. This MLP neural network method can be faster as the neural network model can be saved and reused without necessarily querying the training data on each prediction.

As a summary, our two previous methods can be compared in Table 1.

Table 1 Comparison of the previous methods: VAR and MLP-NN.

	VAR	MLP-NN
Approach	Prediction	Classification
Training Data Components	RSRP/RSRQ, Throughput	RSRP/RSRQ, Download Status
Training Data Usage	Queried from the database on each simulation	Only during the NN model creation (first simulation)
Method for Inference	Query the time series throughput data (filtered with RSRP/RSRQ) from the Training Data and predict the next throughput for each cell	Provide RSRP/RSRQ as input to the NN model and perform classification to obtain successful download probability (score) from each cell

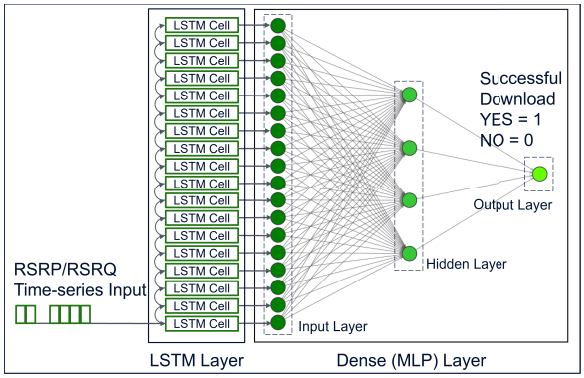


Fig. 7 LSTM neural network design.

6. Proposed Method: LSTM Neural Network

The two previous methods can outperform the traditional handover algorithm. However, the MLP neural network still underperforms the VAR method. We view the neural network as more promising since it has a less complex and less demanding implementation. It requires only RSRP/RSRQ measurements with no throughput measurement and the model can be created once and used repeatedly in every case. Therefore, we seek ways to improve this neural network.

MLP has several limitations and weaknesses: it is stateless, unaware of temporal structure, has messy scaling, and requires fixed-sized inputs and outputs [48]. In our scenario, the input is RSRP/RSRQ measurements which are time-series data, not static fixed-sized data. The length of the input is actually not fixed, depending on the cases. Using MLP, we have to fixate the input to only 3 samples per cell and thus limit the amount of information. Our MLP neural network also has an issue with scaling. When we add or remove the cell amount in the network we have to completely change the model architecture.

In machine learning, time-series data can be handled in various ways. Time-series prediction involves predicting the next value for a given input sequence, time-series classification involves predicting a class label for a given input sequence, and time-series generation involves generating a new output sequence that has the same general characteristics [48].

We assess and identify our scenario as a time-series classification problem because our input data has a temporal structure and we solve the problem by classifying if the file download is successful or not. There are so many methods to solve this time-series classification problem [49], including the deep learning approach using neural network [50]. Instead of using a simple MLP neural network for time-series data, it is recommended to use a Recurrent Neural Network (RNN) that better considers the temporal feature of the input. There are several methods based on RNN, for example, SimpleRNN, Gated Recurrent Unit (GRU), and Long Short-term Memory (LSTM).

Long Short-term Memory (LSTM) [51] is an artificial neural network that has a feedback connection and thus can be classified as RNN. LSTM has been shown to outperform other RNN methods on numerous temporal processing tasks [52]. These temporal processing tasks include the processing of multivariate time-series data to perform predictions on future values. Several applications employ LSTM due to this capability, for example, handwriting recognition, speech recognition, and machine translation.

LSTM employs the "Long-term memory" and "Short-term memory" that occurs in the RNN architecture that processes time-series data. The connection weights and biases in the network change once per episode of training, analogous to how physiological changes in synaptic strengths store long-term memories; the activation patterns in the network change once per time-step, analogous to how the moment-to-moment change in electric firing patterns in the brain store short-term memories [53]. The LSTM architecture aims to provide a short-term memory for RNN that can last thousands of timesteps, thus "Long Short-Term memory". LSTM networks are well-suited to classifying, processing, and making predictions based on time series data since there can be lags of unknown duration between important events in a time series.

We modify our neural network by adding an LSTM layer before the MLP layer. This will enable the neural network to process the time-series input data before feeding them to the MLP to perform classification (Fig. 7).

The LSTM layer contains 18 serially-connected LSTM cells. The input is the RSRP/RSRQ data fed into the first LSTM cell and processed serially to the next LSTM cells. Besides providing input to the next LSTM cell, all 18 LSTM cells also provide input to the MLP's 18 input layers that will further perform the classification function. Using this architecture, any arbitrary length of input can be processed, unlike the MLP neural network that requires fixed-length input. The dropout rate is chosen 50% and the MLP layer architecture is the same with the previous method 18 input nodes, 4 hidden nodes, and 1 output node (Fig. 6). The activation function is *relu* in the hidden layer and *sigmoid* in the output layer. The training process is done with 150 times iterations through the whole training data (*epoch* = *150*), and the model is updated every 10 training data (*batch size* = *10*).

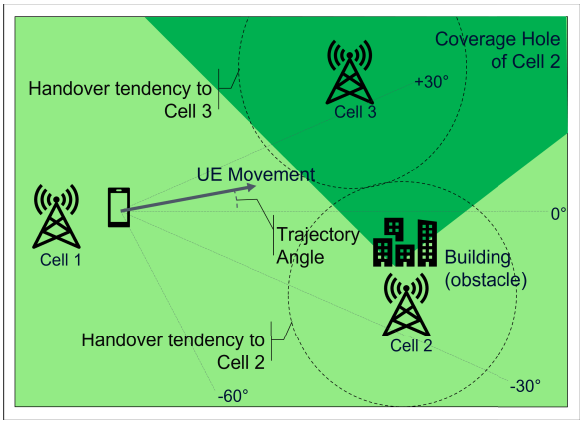


Fig. 8 Environment for simulation with 3 cells and 1 coverage hole.

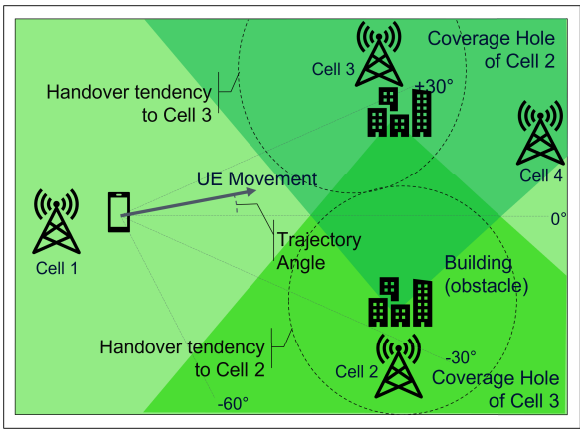


Fig. 9 Environment for simulation with 4 cells and 2 coverage holes.

7. Simulation Design and Data Collection

In this research, we created two network environments for two experiments. The first one contains three cells, one moving UE, and a building creating a coverage hole (Fig. 8). The second one contains four cells, one moving UE, and two buildings creating two coverage holes (Fig. 9). This environment is built using NS3 LTE network simulator [54] based on previous studies[33], [36]. The simulation parameters are reusing the previous work as described in Table 2.

On each simulation, the UE moves to the right side of the network with a random trajectory angle. Due to this movement, the UE needs to perform a handover from Cell 1 to either Cell 2 or Cell 3 (or also Cell 4 for 4 cells simulation), depending on the trajectory angle. The UE also downloads files during the movement and in the end, the download may be successful or may not. For every simulation, we noted down the target cell, the download success status, and the RSRP/RSRQ measured by UE.

The simulation activity can be described in Fig. 10. Our simulation contains three activities: the training data generation (1), handover simulations using the traditional algorithm (2), the target cell determination using Near-RT RIC (3), and the handover result verification of the target cell determined by Near-RT RIC (4).

Table 2 NS3 simulation parameters.

Parameter	Value
System bandwidth	5 MHz
Inter-site distance	500 m
Adaptive Modulation and Coding Scheme	MiErrorModel
Simulation area	2000×2000 m^2
Number of base stations	3 and 4
Transmit Power	46 dBm
Number of UEs	3 and 4 (1 moving)
Velocity of UE1	16.6667 m/s
Path Loss Model	Cost 231
Antenna Height	30 m
Obstacle Height	35 m
File Size	15 MB

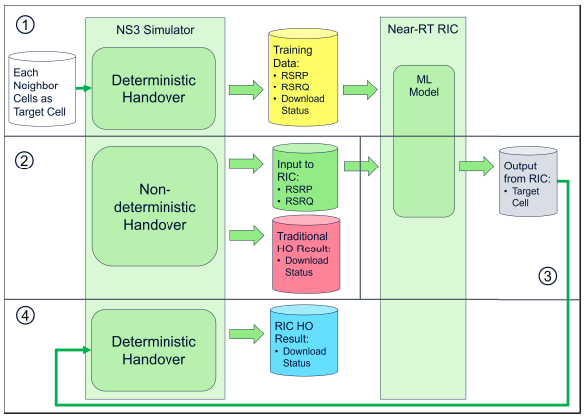


Fig. 10 Simulation activities done for this research.

To create the training data, we ran 100 simulations of deterministic handover to each neighbor cell. The first 100 simulations are deterministic handover cases where the UE is forced to perform handover to Cell 2 regardless of the trajectory angle. The next 100 simulations are also deterministic handover cases but this time to Cell 3. There are also the next 100 simulations to Cell 4 for 4 cells simulation. This activity is described as the first activity in Fig. 10.

To compare the performance of the handover algorithms (the traditional algorithm and our RIC-based proposed algorithms), we ran other simulations of non-deterministic handover. In these simulations, the UE may perform a handover to any neighbor cell, using a traditional handover algorithm, based on the RSRP/RSRQ measurements.The result of these simulations (download success status and RSRP/RSRQ measurement) is the result of the traditional handover algorithm and is used as a baseline to be compared with machine-learning-based algorithms run in Near-RT RIC. The RSRP/RSRQ measurement for these simulations is also used as input for the RIC-based handover algorithm. This activity is described as the second activity in Fig. 10.

Next, we performed RIC-based handover simulations. For each simulation run, we performed target cell determination using the machine-learning-based algorithm in Near-

RT RIC by providing RSRP and RSRQ measurements of the same simulations that we ran in the traditional handover algorithm process. The algorithm in Near-RT RIC would then get the score of each existing neighbor cell in the network. The cell with the highest score is then chosen as the target cell. This activity is described as the third activity in Fig. 10.

From Near-RT RIC we only obtained the target cell, but not yet the download success status. Therefore, we need to perform verification using NS3 to check if the download is successful or not, given the target cell from Near-RT RIC. Next, we performed deterministic handover again using NS3 but using the target cell obtained by Near-RT RIC. From here we get the download success status if the handover is controlled by Near-RT RIC. This activity is described as the fourth activity in Fig. 10.

We choose download success rate as the main performance metric. As stated in Eq. (1) in Sect. 2, the handover process is best improved by reducing the probability of handover failure, thus avoiding unnecessary handover and handover to a wrong cell. Based on this statement, we focus on the target cell determination process. We decide the performance metric as download success rate if we use a certain method to select the target cell.

8. Simulation Result and Discussion

As described in Sect. 7, we already performed three sets of simulations in our previous papers: the traditional handover, Near-RT RIC handover using VAR, and Near-RT RIC using MLP neural network. In this paper, we propose an additional method which is Near-RT RIC using LSTM neural network and we also performed another set of simulations. The traditional handover was done using the NS3 simulator and we noted down the RSRP/RSRQ measurement and the handover results (target cell and download success status). The RSRP/RSRQ measurement of those simulations was used as input in Near-RT RIC handover simulations. After that, we compared the download success rate of all simulations among all methods (Fig. 11 and Fig. 12). We performed all those simulations in two network scenarios: 3 cells with 1 coverage hole and 4 cells with 2 coverage holes.

The successful download rate for the traditional handover algorithm, in 3 cells 1 coverage hole environment, is 86.2%, not 100% due to the presence of the coverage hole. All of the simulations with failed downloads happened when the UE was handed over to Cell 2 (based on the best RSRP/RSRQ measurement) but it entered the coverage hole behind the building after the handover. If it was handed over to Cell 3 instead of Cell 2, the download may be successful because Cell 3 was not obstructed by the building. This result shows that sometimes the traditional handover algorithm is not reliable in a non-ideal condition.

This successful download rate for the traditional handover algorithm is getting worse in the 4 cells 2 coverage holes environment, which reached only 29%. The two coverage holes created a blank spot in the network that caused download failure if the UE performed a handover to either

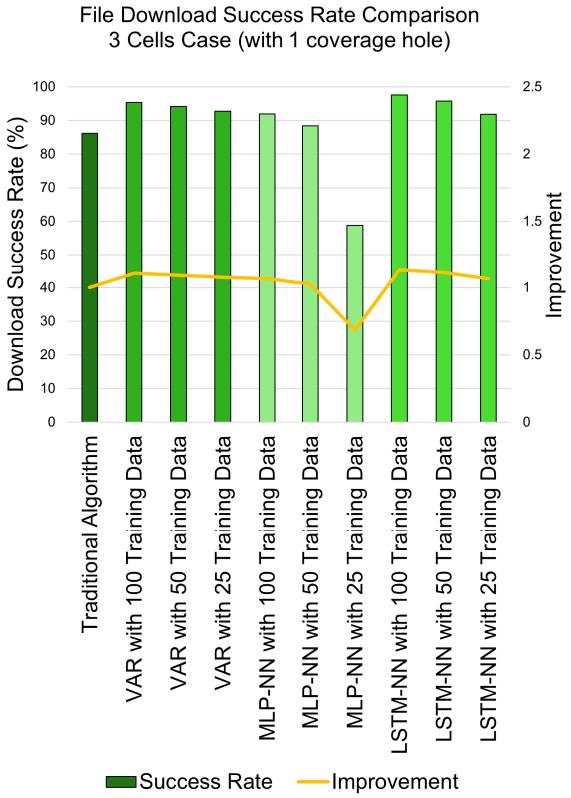


Fig. 11 Simulation result comparison for a network with 3 cells and 1 coverage hole.

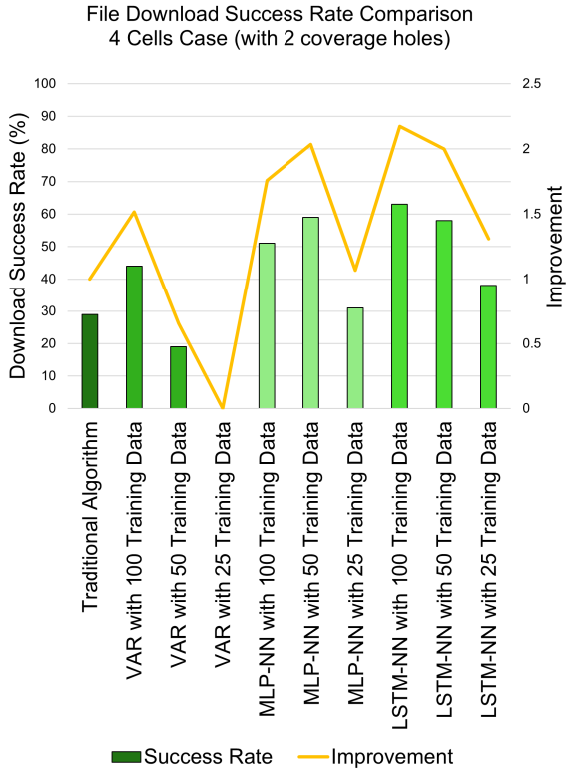


Fig. 12 Simulation result comparison for a network with 4 cells and 2 coverage holes.

Table 3 Comparison of the handover methods.

Method	Traditional	Related Work[6]	Modified VAR	MLP-NN	LSTM-NN
Approach	Selection	Classification	Prediction	Classification	Classification
Brief Description	Select the target cell from the neighbor cell (as the candidate target cell) that has the best RSRP and/or RSRQ.	Provide the RSRP and RSRQ measurement to an MLP neural network model to get the score of a candidate target cell. The score of all candidate cells is then compared and the highest score is chosen as the target cell.	Predict the future data rate of a candidate cell from the pool of training data. The candidate cell with the highest predicted data rate is chosen as the target cell.	Provide the RSRP and RSRQ measurement to an MLP neural network model to get the score of a candidate target cell. The score of all candidate cells is then compared and the highest score is chosen as the target cell	Provide the RSRP and RSRQ measurement to the LSTM network and continue to the MLP model to get the score of a candidate target cell. The score of all candidate cells is then compared and the highest score is chosen as the target cell
Best Performance (download success rate in 1-coverage-hole network)	86.2%	95.37%	95.3%	91.9%	97.6%

Cell 2 or Cell 3. The UE may experience a successful download if it performed a handover to Cell 4 but Cell 4 is never an option in the traditional handover algorithm as the RSRP/RSRQ are too low at the time of handover.

When we determined the target cell using Near-RT RIC in the 3 cells 1 coverage hole network environment, the successful download rates are mostly increasing, depending on the method and the amount of training data (Fig. 11). If the QoE Predictor xApp uses vector autoregression (VAR), the success rate can reach 95.3% using all 100 available training data, 94.1% with 50 training data, and 92.7% with only 25 training data. If we use MLP-NN, the success rate is slightly lower but still higher than the traditional algorithm in most cases, the download success rate can reach 91.9% using all 100 available training data and 88.4% using only 50 training data. However, the performance plummeted to only 58.8% if we only use 25 training data (even lower than the traditional handover algorithm). Using our newly-proposed LSTM-NN, the success rate is superior to other methods, reaching 97.6% using 100 training data, 95.7% using 50 training data, and 91.8% using only 25 training data.

If we use Near-RT RIC to determine the target cell in the 4 cells 2 coverage holes network environment, some methods significantly can increase the success rate, given enough training data. As shown in Fig. 12, VAR didn't perform much in this network environment. Our previous MLP-NN method performed better but ultimately our newly-proposed LSTM-NN worked best for this environment.

The related work [6] reported that their proposed method achieved a download success rate of 95.37% compared to the state-of-the-art method that yields only 54.45% in their simulation. However, they did not perform the implementation in the Near-RT RIC.

It is shown that machine-learning algorithms can provide better handover performance by determining the correct target cell in a non-ideal network condition, given enough training data. The more training data, the better the performance. From the simulation result, our newly proposed LSTM-NN works better than all other methods and we proved it in both network environment cases.

As a summary, all of the handover methods discussed in this paper can be compared in Table 3. The Modified VAR and MLP-NN are our previous proposed methods [13], [14] while LSTM-NN is the current proposed method.

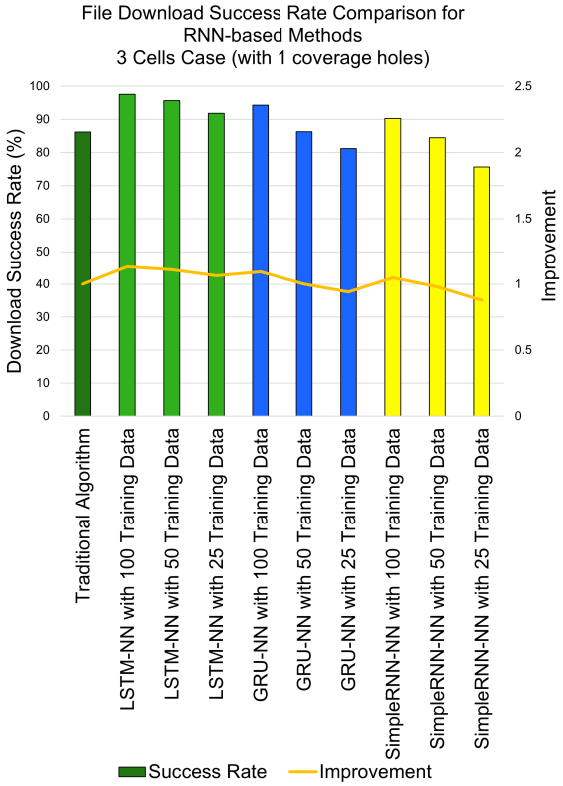


Fig. 13 Comparison of RNN-based methods for a network with 3 cells and 1 coverage hole.

In addition, we also compared other RNN-based methods that are designed for time-series classification. We replaced the LSTM layer with the GRU and SimpleRNN layer. No changes were performed in the MLP layer and the hyperparameters. The result is similar to the LSTM method but LSTM is still superior to those methods (Fig. 13). However, those methods are very promising for future exploration.

9. Conclusion and Future Work

The handover process may cause an interruption in the data transmission, moreover in a high-mobility condition where the radio condition may worsen because of the user speed. Increasing the probability of successful handovers, such as making sure to perform handover to the correct target cell, can minimize this interruption. Therefore, target cell determination is very important in the handover process.

In this paper, we presented the result of our newly proposed method, the LSTM neural network, using O-RAN Near-RT RIC to determine the target cell in the handover process. This new method is an improvement of our previous machine-learning-based methods that is better at handling the time-series nature of the input data. From the simulation result, it can be concluded that this method can be used and is proven better to determine the target cell compared to other methods, the traditional handover algorithm, and our previous machine-learning-based methods. The performance of the algorithms depends on the method and the amount of training data.

In the future, we will further improve the neural network to get better performance. We plan to test this method in another non-ideal network environment other than the coverage hole case. We also plan to explore the possibilities of using another RNN-based methods such as GRU and SimpleRNN.

References

[1] M.S. Obaidat and P. Nicopolitidis, Smart Cities and Homes: Key Enabling Technologies, Morgan Kaufmann, 2016.
[2] H.S. Park, Y. Lee, T.J. Kim, B.C. Kim, and J.Y. Lee, "Handover mechanism in NR for ultra-reliable low-latency communications," IEEE Netw., vol.32, no.2, pp.41–47, 2018.
[3] 3GPP, "NR; User Equipment (UE) radio transmission and reception; part 2: Range 2 standalone-Rel. 16," 2020.
[4] P.V. Klaine, M.A. Imran, O. Onireti, and R.D. Souza, "A survey of machine learning techniques applied to self-organizing cellular networks," IEEE Commun. Surveys Tuts., vol.19, no.4, pp.2392–2431, 2017.
[5] C. Lee, H. Cho, S. Song, and J.M. Chung, "Prediction-based conditional handover for 5G mm-Wave networks: A deep-learning approach," IEEE Veh. Technol. Mag., vol.15, no.1, pp.54–62, 2020.
[6] Z. Ali, N. Baldo, J. Mangues-Bafalluy, and L. Giupponi, "Machine learning based handover management for improved QoE in LTE," Proc. NOMS 2016 - 2016 IEEE/IFIP Network Operations and Management Symposium, no.5, pp.794–798, 2016.
[7] N. Sinclair, D. Harle, I.A. Glover, J. Irvine, and R.C. Atkinson, "An advanced SOM algorithm applied to handover management within LTE," IEEE Trans. Veh. Technol., vol.62, no.5, pp.1883–1894, 2013.
[8] N.M. Alotaibi and S.S. Alwakeel, "A neural network based handover management strategy for heterogeneous networks," Proc. 2015 IEEE 14th International Conference on Machine Learning and Applications, ICMLA 2015, pp.1210–1214, 2016.
[9] M.A.F. Rihani, M. Mroue, J.C. Prevotct, F. Nouvel, and Y. Mohanna, "A neural network based handover for multi-RAT heterogeneous networks with learning agent," Proc. 13th International Symposium on Reconfigurable Communication-Centric Systems-on-Chip, ReCoSoC 2018, 2018.
[10] B. Shubyn, N. Lutsiv, O. Syrotynskyi, and R. Kolodii, "Deep learning based adaptive handover optimization for ultra-dense 5G mobile networks," Proc. 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering, TCSET 2020, pp.869–872, 2020.
[11] G. Masini, Y. Gao, and S. Sirotkin, "Artificial intelligence and machine learning in NG-RAN: New study in RAN3," 2021.
[12] O-RAN, "O-RAN: Towards an Open and Smart RAN," White Paper, Oct. 2018.
[13] B.H. Prananto, Iskandar, and A. Kurniawan, "Handover improvement using O-RAN near-RT RIC for network with coverage hole," IEICE Commun. Express, vol.11, no.1, pp.19–23, 2023.
[14] B.H. Prananto, Iskandar, and A. Kurniawan, "A new method to improve frequent-handover problem in high-mobility communications using RIC and machine learning," IEEE Access, vol.11, pp.72281–72294, 2023.
[15] J. Agrawal, P. Mor, J. Keller, R. Patel, and P. Dubey, "Introduction to the basic LTE handover procedures," 2015 International Conference on Communication Networks (ICCN), pp.197–201, 2015.
[16] 3GPP TS 36.300, "3GPP TS 36.300 V16.0.0 (2019-12) evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN) overall description Stage 2 (Release 10)," 2013.
[17] 3GPP TR 38.913, "Study on scenarios and requirements for next generation access technologies (Release 16)," 2020.
[18] 3GPP TR 36.881, "Evolved universal terrestrial radio access (E-UTRA) study on latency reduction techniques for LTE (Release 14)," 2016.
[19] A. Chincholi, M. Menon, and L. Hsu, "Measurement gap enhancements for BL/CE UEs," World Intellectual Property Organization WO 2019/094977 Al, May 2019.
[20] M. Polese, M. Giordani, M. Mezzavilla, S. Rangan, and M. Zorzi, "Improved handover through dual connectivity in 5G mmWave mobile networks," IEEE J. Sel. Areas Commun., vol.35, no.9, pp.2069–2084, 2017.
[21] H.L. Wang, S.J. Kao, C.Y. Hsiao, and F.M. Chang, "A moving direction prediction-assisted handover scheme in LTE networks," J. Wireless Com. Network., vol.2014, no.1, 190, 2014.
[22] H. Martikainen, I. Viering, A. Lobinger, and T. Jokela, "On the basics of conditional handover for 5G mobility," IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC, vol.2018-Septe, no.2, 2018.
[23] I.L. Da Silva, C. Eklöf, J. Muller, and R. Zhohov, "This is the key to mobility robustness in 5G networks," Ericsson.com, https://www.ericsson.com/en/blog/2020/5/the-key-to-mobility-robustness-5g-networks, accessed March 13 2024.
[24] Y. Koda, K. Yamamoto, T. Nishio, and M. Morikura, "Reinforcement learning based predictive handover for pedestrian-aware mmWave networks," INFOCOM 2018 - IEEE Conference on Computer Communications Workshops, pp.692–697, 2018.
[25] C. Wang, L. Ma, R. Li, T.S. Durrani, and H. Zhang, "Exploring trajectory prediction through machine learning methods," IEEE Access, vol.7, pp.101441–101452, 2019.
[26] R.V. Akhpashev and V.G. Drozdova, "Spatial interpolation of LTE measurements for minimization of drive tests," International Conference of Young Specialists on Micro/Nanotechnologies and Electron Devices, EDM, vol.2018-July, pp.136–138, 2018.
[27] A. Suresh Kumar, S. Vanmathi, B. Praveen Sanjay, S. Ramya Bharathi, and M. Sakthi Meena, "Handover forecasting in 5G using machine learning," International Journal of Engineering and Technology (UAE), vol.7, no.2, pp.76–79, 2018.
[28] A. Masri, T. Veijalainen, H. Martikainen, S. Mwanje, and J. Ali-Tolppa, "Machine-learning-based predictive handover," IEEE IM 2021- France, no.Im, pp.648–652, 2021.
[29] Expert.AI Team, "What is machine learning?," Expert.AI, https://www.expert.ai/blog/machine-learning-definition/, accessed March 13 2024.
[30] T. Mitchell, Machine Learning, McGraw Hill, 1997.

[31] X. Chen, F. Meriaux, and S. Valentin, "Predicting a user's next cell with supervised learning based on channel states," IEEE Workshop on Signal Processing Advances in Wireless Communications, SPAWC, pp.36–40, 2013.
[32] L. Yan, H. DIng, L. Zhang, J. Liu, X. Fang, Y. Fang, M. Xiao, and X. Huang, "Machine learning-based handovers for sub-6 GHz and mmWave integrated vehicular networks," IEEE Trans. Wireless Commun., vol.18, no.10, pp.4873–4885, 2019.
[33] T. Cabral De Brito, G. Advisor, and V.A. De Sousa, "Machine learning based handover management for LTE networks with coverage holes," Ph.D. Thesis, Universidade Federal do Rio Grande do Norte, 2018.
[34] L.L. Vy, L.P. Tung, and B.S.P. Lin, "Big data and machine learning driven handover management and forecasting," 2017 IEEE Conference on Standards for Communications and Networking, CSCN 2017, no.September 2019, pp.214–219, 2017.
[35] V. Yajnanarayana, H. Ryden, and L. Hevizi, "5G handover using reinforcement learning," 2020 IEEE 3rd 5G World Forum, 5GWF 2020 - Conference Proceedings, pp.349–354, 2020.
[36] Z. Ali, N. Baldo, J. Mangues-Bafalluy, and L. Giupponi, "Simulating LTE mobility management in presence of coverage holes with ns-3," SIMUTOOLS 2015 - 8th EAI International Conference on Simulation Tools and Techniques, 2015.
[37] J. Nakazato, M. Kuchitsu, A. Pawar, S. Masuko, K. Tokugawa, K. Kubota, K. Maruta, and K. Sakaguchi, "Proof-of-concept of distributed optimization of micro-services on edge computing for beyond 5G," IEEE Vehicular Technology Conference, vol.2022-June, pp.1–6, 2022.
[38] A. Akman, "O-RAN Working Group 1 Use Cases Detailed Specification," O-RAN Alliance e.V. Technical Specification, 2020.
[39] Nokia, "Nokia launches first commercial Service Enablement Platform to drive Open RAN innovation," Nokia Press Release, https://www.nokia.com/about-us/news/releases/2021/03/11/nokia-launches-first-commercial-service-enablement-platform-to-drive-open-ran-innovation/, accessed March 13 2024.
[40] O. Orhan, V.N. Swamy, T. Tetzlaff, M. Nassar, H. Nikopour, and S. Talwar, "Connection management xAPP for O-RAN RIC: A graph neural network and reinforcement learning approach," 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA), pp.936–941, 2021.
[41] Y. Li, E. Datta, J. Ding, N.B. Shroff, and X. Liu, "Can online learning increase the reliability of extreme mobility management?," 2021 IEEE/ACM 29th International Symposium on Quality of Service, IWQOS 2021, pp.0–5, 2021.
[42] L. Bonati, S. D'Oro, M. Polese, S. Basagni, and T. Melodia, "Intelligence and learning in O-RAN for data-driven NextG cellular networks," IEEE Commun. Mag., vol.59, no.10, pp.21–27, 2021.
[43] H. Lee, Y. Jang, J. Song, and H. Yeon, "O-RAN AI/ML workflow implementation of personalized network optimization via reinforcement learning," 2021 IEEE Globecom Workshops (GC Wkshps), pp.1–6, 2021.
[44] F. Cefalu, "O-RAN SC Wiki: Getting started," O-RAN SC Confluence, https://wiki.o-ran-sc.org/display/GS/Getting+Started, accessed March 13 2024.
[45] D. Karnwal, "Anomaly detection use case," O-RAN SC Confluence, https://wiki.o-ran-sc.org/display/RICP/Anomaly+Detection+Use+Case, accessed March 13 2024.
[46] B.H. Prananto, Iskandar, and A. Kurniawan, "Study on neural network for cellular mobility management: NS3 simulation with coverage hole case," Proc. 15th International Conference on Telecommunication Systems, Services, and Applications, TSSA 2021, pp.1–6, 2021.
[47] B.H. Prananto, Iskandar, and A. Kurniawan, "O-RAN intelligent application for cellular mobility management," 9th International Conference on ICT for Smart Society: Recover Together, Recover Stronger and Smarter Smartization, Governance and Collaboration, ICISS 2022 - Proceeding, pp.1–6, 2022.
[48] J. Brownlee, "Long short-term memory networks with Python: Develop sequence prediction models with deep learning," 2017.
[49] J. Faouzi, "Time series classification: A review of algorithms and implementations," Machine Learning (Emerging Trends and Applications), hal-03558165, in press.
[50] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.A. Muller, "Deep learning for time series classification: A review," Data Min. Knowl. Disc., vol.33, no.4, pp.917–963, 2019.
[51] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol.9, no.8, pp.1735–1780, 1997.
[52] F.A. Gers, D. Eck, and J. Schmidhuber, "Applying LSTM to time series predictable through time-window approaches," Artificial Neural Networks — ICANN 2001, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol.2130, pp.669–676, 2001.
[53] J.L. Elman, "Finding structure in time," Cognitive Science, vol.14, no.2, pp.179–211, 1990.
[54] NS3, "Design documentation — Model library."

Baud Haryo Prananto received a B.S. degree from Bandung Institute and Technology, Indonesia in 2004 and an M.S. degree from Korea Institute of Science and Technology, South Korea in 2008. Currently, he is studying at Bandung Institute of Technology as a Doctoral student. He is currently working as an Expert Technical Trainer in Nokia Solutions and Networks since 2008, delivering training related to 4G and 5G RAN Nokia equipment.

Iskandar received B.S. and M.S. degrees from Bandung Institute and Technology, Indonesia in 1995 and 2000 respectively, and a Doctoral degree from Waseda University Tokyo, Japan in 2007. Currently, he is an Associate Professor in the School of Electrical Engineering and Informatics, Bandung Institute of Technology with research interests mainly in wireless telecommunications.

Hendrawan received a B.S. degree from Bandung Institute and Technology, Indonesia in 1985 and an M.S. and Ph.D. degree from the University of Essex, UK in 1990 and 1995 respectively. Currently, he is an Associate Professor in the School of Electrical Engineering and Informatics, Bandung Institute of Technology with research interests mainly in queuing theory, data communication, multimedia communication, telecommunication network, and network management.

Adit Kurniawan received a B.S. degree from Bandung Institute and Technology, Indonesia in 1986, an M.S. degree from Royal Melbourne Institute of Technology, Australia in 1996, and a Ph.D. degree from the University of South Australia in 2003. He was a Professor in the School of Electrical Engineering and Informatics, Bandung Institute of Technology with research interests in antenna propagation, microwave propagation, wireless spread spectrum, and CDMA.

IEICE TRANS. COMMUN., VOL.E107–B, NO.6 JUNE 2024

PAPER

A Novel Remote-Tracking Heart Rate Measurement Method Based on Stepping Motor and mm-Wave FMCW Radar

Yaokun HU †,††a), *Student Member*, **Xuanyu PENG** †††*, *and* **Takeshi TODA** ††††**, *Members*

SUMMARY The subject must be motionless for conventional radar-based non-contact vital signs measurements. Additionally, the measurement range is limited by the design of the radar module itself. Although the accuracy of measurements has been improving, the prospects for their application could have been faster to develop. This paper proposed a novel radar-based adaptive tracking method for measuring the heart rate of the moving monitored person. The radar module is fixed on a circular plate and driven by stepping motors to rotate it. In order to protect the user's privacy, the method uses radar signal processing to detect the subject's position to control a stepping motor that adjusts the radar's measurement range. The results of the fixed-route experiments revealed that when the subject was moving at a speed of 0.5 m/s, the mean values of RMSE for heart rate measurements were all below 2.85 beat per minute (bpm), and when moving at a speed of 1 m/s, they were all below 4.05 bpm. When subjects walked at random routes and speeds, the RMSE of the measurements were all below 6.85 bpm, with a mean value of 4.35 bpm. The average RR interval time of the reconstructed heartbeat signal was highly correlated with the electrocardiography (ECG) data, with a correlation coefficient of 0.9905. In addition, this study not only evaluated the potential effect of arm swing (more normal walking motion) on heart rate measurement but also demonstrated the ability of the proposed method to measure heart rate in a multiple-people scenario.

key words: *FMCW radar, health care, heart rate, radar signal processing, vital sign detection*

1. Introduction

The number of seniors living alone and in nursing homes is rising because of longevity. It is difficult to detect potential medical problems in the body of a senior person. However, cardiovascular disease is the primary cause of death in Japan, accounting for 25.5% of all deaths [1]. The development of an indoor heart rate monitoring system for older people living alone is urgently needed because of this growing social issue.

The conventional contact measures are inappropriate for continuous, round-the-clock monitoring and seniors with skin issues. The research on non-contact radar-based heart rate measurements has recently gained popularity. In [2], they estimated the subject's vital signs were estimated using a millimeter-wave FMCW radar. The authors of [3] proposed a differential enhancement approach and employed an FMCW radar with an 8.4 GHz center carrier frequency to precisely measure the heart rate. In [4], a learning-refined integral null space pursuit algorithm was suggested for adaptively acquiring vital signs with an impulse radio UWB (IR-UWB) radar and a stepped-frequency continuous-wave UWB radar. For non-contact vital sign monitoring, a continuous wave (CW) Doppler radar that operates as a phase-locked loop in a phase-demodulator configuration was proposed in [5].

IR-UWB radar [4], [8], [9], CW Doppler radar [5], [10]–[14], and FMCW radar [2], [15]–[18] are the types of radars that are currently most frequently utilized in this study field. The maximum signal energy levels that can be transmitted are not very high, which reduces the precision and signal-to-noise ratio (SNR) of IR-UWB radars [19]. Likewise, because CW Doppler radar lacks a range capacity [20], measurements are subject to interference and are therefore inappropriate for keeping track of the vital signs of moving targets. The FMCW radar combines the advantages of the first two, with good range and speed measurement capabilities, and the millimeter-band FMCW radar is extremely sensitive. It can effectively measure the displacement of the human skin surface. Therefore, this study employs the FMCW radar in the 77–81 GHz band.

Most radar-based vital signs estimation studies call for subjects to stay still, including standing, sitting in a chair, or resting in bed [2]–[20]. This limits the application of radar-based heart rate measurement methods. For this reason, the authors of [21], [22] and [23] attempted to use IR-UWB radar and FMCW radar, respectively, to estimate the heart rate of a moving person. However, the subjects in the research mentioned above could only move slowly. To address this issue, we proposed an improved adaptive range bin selection (IARBS) method for moving subjects' heart rate measurements based on the FMCW radar [24]. The subjects maintained a high measurement accuracy when they walked at an average speed of 1 m/s. However, because of the radar characteristics, the measurement range may be constrained. It is challenging to cover the entire room, even with beamforming techniques. When the angle is too large, the beam widens, the antenna gain drops, and the perfor-

Manuscript received August 23, 2023.
Manuscript revised October 25, 2023.
Manuscript publicized January 30, 2024.

†The author is with the Graduate School of Science and Technology, Nihon University, Tokyo, 101-0062 Japan.
††The author is with Fujitsu Ltd., Kawasaki-shi, 211-8588 Japan.
†††The author was with the College of Intelligent Manufacturing, Hunan Vocational Institute of Technology, Hunan, 411104 China.
††††The author was with Fujitsu Laboratories Ltd., Kawasaki-shi, 211-8588 Japan.
*Presently, with the Graduate School of Science and Technology, Nihon University.
**Presently, with the College of Science and Technology of Nihon University.
a) E-mail: csgy19014@g.nihon-u.ac.jp
DOI: 10.23919/transcom.2023EBP3143

mance suffers. If there are several radars in the room, they can interfere with each other and increase background noise. Moreover, these methods increase the cost of the measurement system, which is not conducive to the popularity of the application.

The authors of [34] proposed a novel approach for mounting radar on a mobile robot. However, it may not be appropriate for older adults living alone, in senior homes, or in medically isolated observations. There is a safety risk because the senior may trip over the wires because the robot must be attached to the power source. It is difficult to continuously monitor if a charging method is employed. In [26], a method for adaptively changing radar orientation by acquiring the subject's position through a camera was proposed. However, image processing using cameras is sensitive to ambient lighting and involves personal privacy issues. Additionally, the subjects' heart rate measurements while moving were not evaluated in the abovementioned study.

In [27], we proposed and discussed the possibility of using a stepping motor to change the radar orientation but only verified it using a simple simulation. This paper improves the method of [27] and proposes a novel, radar-based, adaptive remote-tracking heart rate measuring technique to address the abovementioned issues. In this study, we build an actual measurement device. The radar module is fixed to a rotatable circular plate. The radar measures the heart rate while measuring the position of the monitored person, which is then processed and transmitted to a stepping motor to drive the rotation of the radar module. A stepping motor is used in this study because for the following reasons. Because the stepping motor has no brushes, it is more reliable. Additionally, it does not accumulate the error of one step to the next step, and thus has better position accuracy. Furthermore, the motor is simple and inexpensive, because its response is based only on a digital input pulse.

The monitored person can always be within the ideal measurement range owing to the adaptively adjusted radar orientation. Next, after sampling the raw radar data, the range bins where the subject is located are selected along the time dimension using the IARBS method, and their phase information is extracted. The displacement information of the human skin surface, brought on by the body's torso, heartbeat, and breathing, can be obtained using the phase change information. We suggested using the improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) approach to extract the heartbeat signal and determine its heart rate [28]. In this study, we continue to use this technique to analyze the experimental data and evaluate the accuracy of the proposed measurement system.

In this study, we build an actual measurement device, using a stepping motor to change the radar direction to adjust the ideal range for measuring the heart rate. The proposed measurement system tracks a moving subject using radar signal processing. Even though this subject walks at an average speed of 1 m/s and the horizontal angle to the radar module varies widely. This significantly increases the application potential of radar-based remote heart rate measurements. To the best of our knowledge, this is the first study that challenges this field.

The principle of employing FMCW radar to detect vital signs is explained in Sect. 2. The proposed method is explained in detail in Sect. 3. The experiments and conclusions are presented in Sects. 4 and 5, respectively.

2. Principle for Heart Rate Measurements

Figure 1 depicts the structure of the FMCW radar-based heart rate measuring system for moving people. As shown in Fig. 2, the ramp generator periodically outputs an up-chirp signal with a time duration of T_c and frequencies ranging from f_{min} to f_{max} at specific time intervals. The slope of the upward linear frequency modulation of each chirp is K_s, and the sweeping bandwidth is $B = f_{max} - f_{min=}T_cK_s$.

The power amplifier amplifies each chirp signal before transmission. The radar module receives the signal after it is reflected by the object and amplifies it using a low-noise amplifier. The transmitted and the received signals are correlated by the in-phase and quadrature (I/Q) mixer to generate the intermediate frequency (IF) signal. Since there are four receivers (with the same configuration), there are also four IF signals. After the analog-to-digital converter (ADC) samples each IF signal, phase information can be preserved. The sampled raw data are transmitted frame-by-frame to the computer for signal processing. Each IF signal can be defined by the following equation:

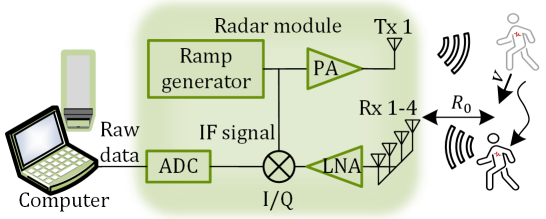


Fig. 1 The block diagram of the FMCW radar module. The analog-to-digital converter, low-noise amplifier, and power amplifier are each denoted by the acronyms PA, LNA, and ADC. The computer receives the raw data to process the signals.

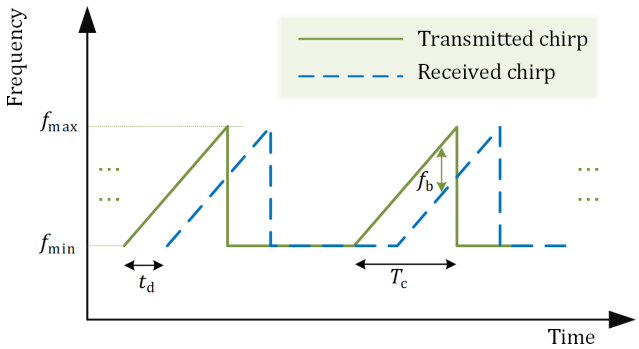


Fig. 2 The time-frequency diagram of the transmitted and received chirps of the case of single Tx and single Rx.

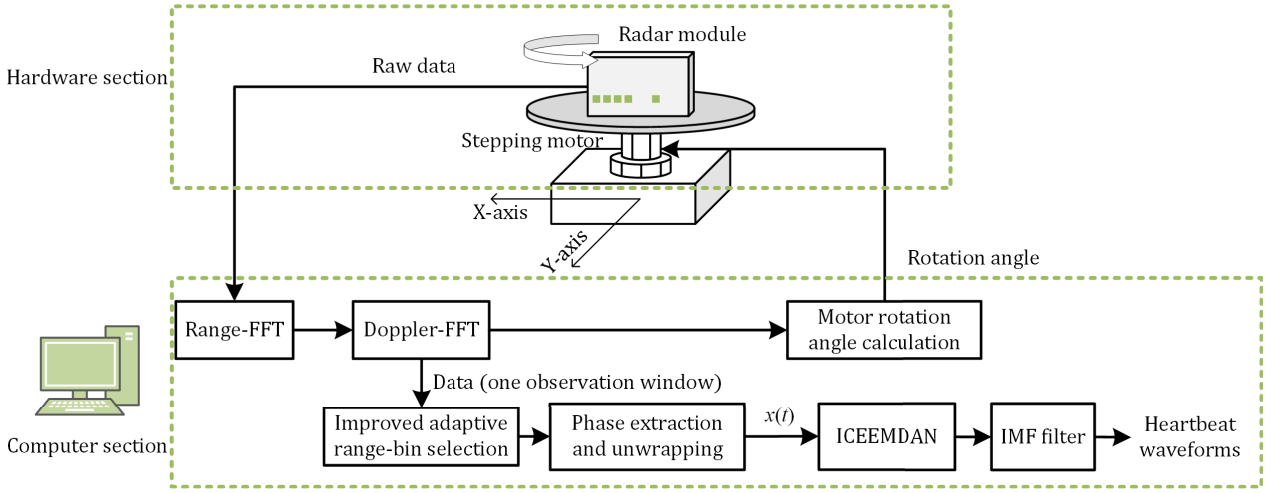


Fig. 3 The flow chart of the proposed processing.

$$s_{\mathrm{if}}(t) = A_{\mathrm{t}}A_{\mathrm{r}} \exp\left(j\left(2\pi f_{\min} t_{\mathrm{d}} + 2\pi K_{\mathrm{s}} t_{\mathrm{d}} t - \pi K_{\mathrm{s}} t_{\mathrm{d}}^2\right)\right)$$
$$\approx A_{\mathrm{t}}A_{\mathrm{r}} \exp\left(j\left(2\pi f_{\min} t_{\mathrm{d}} + 2\pi K_{\mathrm{s}} t_{\mathrm{d}} t\right)\right) t_{\mathrm{d}} < t < T_{\mathrm{c}} \quad (1)$$

where A_{t} is magnitude associated with the transmission power and A_{r} is related to A_{t} by the radar equation. Time delay, t_{d} between the signal being transmitted and received. Because $\pi K_{\mathrm{s}} t_{\mathrm{d}}^2$ in (1) is small, it can be disregarded [2], [28]. The relationship between the instantaneous distance R_0 and delay time t_{d} of the measured target and radar module can be expressed as:

$$R_0 = \frac{\mathrm{c}t_{\mathrm{d}}}{2} \quad (2)$$

where c represents the speed of light. Combining (1) and (2), the frequency of the IF signal and its phase information can be expressed as

$$f_{\mathrm{b}} = \frac{2K_{\mathrm{s}}R_0}{\mathrm{c}}, \quad \varphi(t) = 4\pi f_{\min}\frac{(R_0 + x(t))}{\mathrm{c}} \quad (3)$$

where $x(t)$ can be interpreted as the displacement of the thoracic cavity caused by the heartbeat and breathing of the target. The heart rate estimation can usually be accomplished with only one IF signal.

3. Proposed Measurement System

As mentioned in the introduction, the inability to alter the radar module's orientation is a drawback of conventional radar-based heart rate measurement methods. Therefore, it is impossible to estimate the heart rate once the subject has departed from a specific range. A heart rate measurement method with a motor combined with a millimeter-wave FMCW radar is proposed in this study as a solution to the abovementioned issue and to increase the applicability of the measurement method. The proposed approach utilizes radar signal processing to calculate the target's motion information, which is then sent to a stepping motor to adjust the orientation of the radar module. In this manner, real-time adaptive tracking heart rate monitoring is accomplished.

The signal processing flowchart of the proposed method is shown in Fig. 3, where the computer section is divided into sections for heart rate monitoring and stepping motor rotation control.

3.1 Stepping Motor Rotation Control

The ADC samples the IF signal, which is then sent to a computer in frames for signal processing. In radar-based heart rate measurement research, 0.05 s is a typical choice for the frame period [2], [15]. A frame period T_{f} of 0.05 s is comparable to sampling the target's heartbeat signal at a sampling frequency of 20 Hz because the average human heart rate is less than 2 Hz, which satisfies the sampling theorem. However, excessive shortening of the frame period leads to more frequent data transmission from the radar module to the computer, which may increase the possibility of data loss. Therefore, in this study, the frame length was fixed at 0.05 s.

As shown in Fig. 4, this study emits L chirp signals in a single frame cycle to execute multidimensional FFT calculations, in contrast to other studies that emit only one chirp signal in each frame. A raw data matrix with L rows and N columns is generated for each frame after each chirp signal is sampled N times. A range-FFT is first performed to the fast time dimension of each frame's raw data to generate the range profile matrix (RPM). Next, a Doppler-FFT is applied to the slow time dimension of the RPM to create a range-Doppler matrix (RDM).

A flowchart of the calculation of the rotation angle of the stepping motor is shown in Fig. 5. Assuming a frame period of 0.05 s, the maximum distance moved in a single frame is only 10 cm, even if the subject walks at a higher speed of 2 m/s. Therefore, the distance moved by the subject in a single frame cycle is relatively small, and considering the high real-time requirements of this measurement method, it is not necessary to rotate the motor frame by

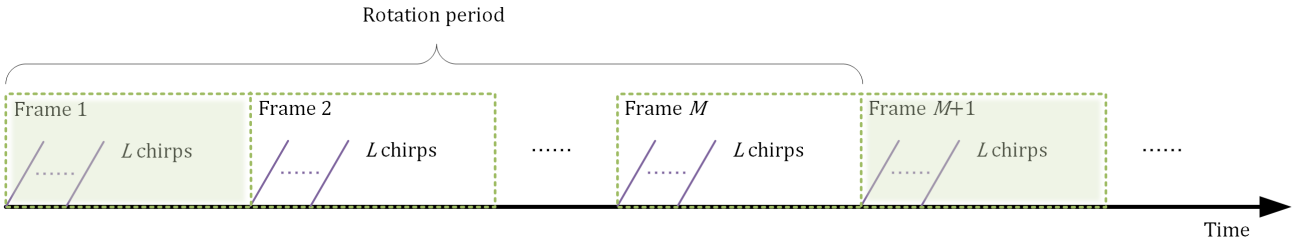


Fig. 4 Diagram of frame period in relation to rotation period.

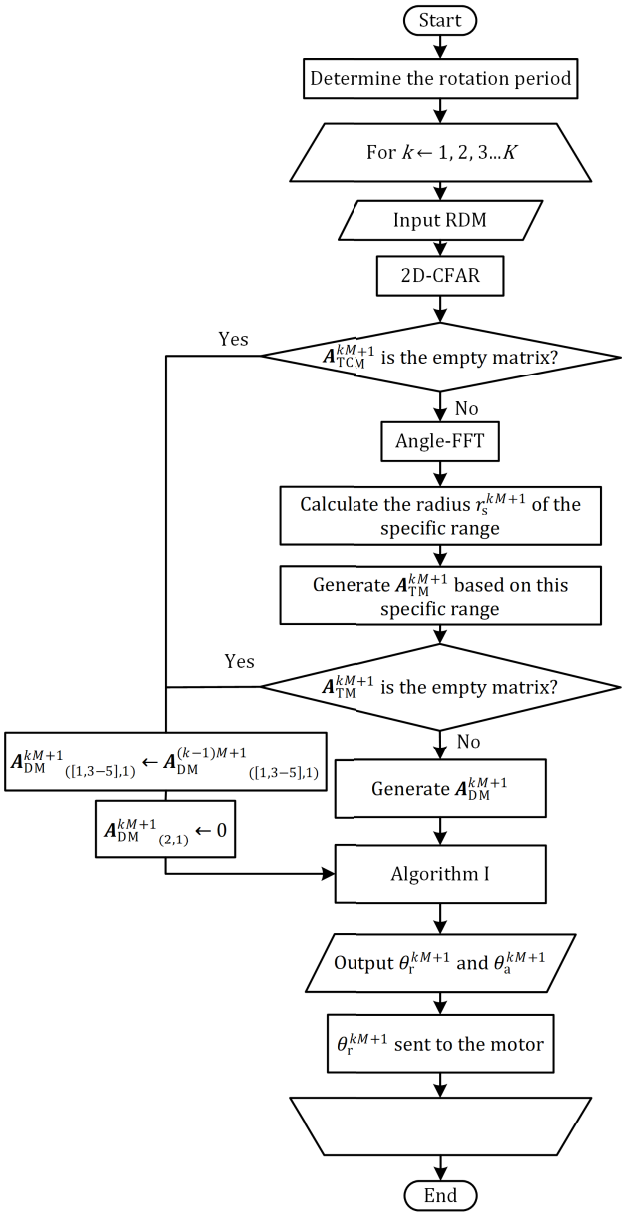


Fig. 5 Flowchart of calculating the rotation angle.

frame. Instead, it determines whether a stepping motor is required to correct the radar alignment by detecting the subject's position every M frames (one rotation period T_r). The motor rotation angle calculation module receives the RDM of the $kM + 1$ th frame ($k = 1, 2, 3, \ldots, K$) as input, as depicted in Fig. 4.

The clutter brought on by other objects and the noise produced by the radar receiver due to temperature, weather, and other factors influence the quality of the radar signal. Therefore, the range bin cells containing the target candidates and their Doppler shifts are identified by performing 2D-CA-CFAR processing for each RDM. 2D-CA-CFAR, 2D-OS-CFAR, and other CFAR algorithms are increasingly being utilized [29], [30]. 2D-OS-CFAR performs better in scenarios with numerous strong interference targets, but the computational effort is substantial [31]. 2D-CA-CFAR is computationally simple and exhibits good performance in pure-noise situations. The 2D-CA-CFAR algorithm is chosen because the environmental background of this study is not complex, the subject is only one person, and the requirement for real-time performance is high. In addition, because the millimeter wave has a high sweep bandwidth, the range bin cell corresponds to a tiny distance interval; therefore, the protection and reference cells of the CA-CFAR algorithm do not affect the measurable range of the radar. CFAR-related algorithms are not the main topic of this work, and the correlations between false alarm probability, judgment threshold, and detection probability are discussed in more detail in [31].

Objects in the RDM that exceed the threshold after 2D-CA-CFAR processing are referred to as target candidates. The instantaneous velocities of the target candidates in the $kM + 1$ th frame are calculated using the Doppler shift. All target candidates, including the subject, remain immobile when the velocities are zero. As a result, the rotation angle θ_r^{kM+1} is zero, and the motor does not need to be adjusted to change the orientation of the radar module.

A target candidate matrix (TCM) $\boldsymbol{A}_{\mathrm{TCM}}^{kM+1}$ is created if u ($u = 1, 2, 3, \ldots, U$) moving targets are detected in the $kM + 1$th frame, with the number of columns representing the number of candidates. The first row is the distance R_{kM+1} from the radar, and the second row represents the instantaneous velocity information v_{kM+1}, as shown in (4).

$$\boldsymbol{A}_{\mathrm{TCM}}^{kM+1} = \begin{bmatrix} R_1^{kM+1} & R_2^{kM+1} & R_3^{kM+1} & \cdots & R_U^{kM+1} \\ v_1^{kM+1} & v_2^{kM+1} & v_3^{kM+1} & \cdots & v_U^{kM+1} \end{bmatrix} \tag{4}$$

The instantaneous angle of the object can be calculated by the following

$$\theta = \sin^{-1} \frac{\lambda \Delta\omega}{2\pi \mathrm{d}}, \tag{5}$$

where λ is the wavelength, $\Delta\omega$ is the phase difference obtained after angle-FFT processing in the antenna direction based on the data obtained by RDM, and d is the pitch of the receiver antennas. The geometric significance of θ is the angle of arrival (AoA) of the object with the radar module in the horizontal plane.

It is assumed that the radar coordinate point is the origin of the coordinate system, and the initial orientation of the radar module is the positive direction of the y-axis, from which the absolute coordinate system is established. The initial position and velocity of the target are $(0, R_1^0)$ and zero, respectively. Additionally, angle $\theta_{\mathrm{a}}^{kM+1}$ is formed by the initial and current radar orientations and is updated in real time. The orientation of the radar changes when tracking the target. Therefore, utilizing $\theta_{\mathrm{a}}^{kM+1}$ to map the target's current coordinates into the initial absolute coordinates is necessary. In this way, the TCM can be expanded to five rows, as shown in (6), with the third row representing the angle information and the fourth and fifth rows representing the x-axis and y-axis absolute coordinates, respectively. The x-axis and y-axis directions are defined in Fig. 3.

$$\boldsymbol{A}_{\mathrm{TCM}}^{0} = \begin{bmatrix} R_1^0 \\ 0 \\ 0 \\ 0 \\ R_1^0 \end{bmatrix},$$

$$\boldsymbol{A}_{\mathrm{TCM}}^{kM+1} = \begin{bmatrix} R_1^{kM+1} & \cdots & R_U^{kM+1} \\ v_1^{kM+1} & \cdots & v_U^{kM+1} \\ \theta_1^{kM+1} & \cdots & \theta_U^{kM+1} \\ R_1^{kM+1}\sin(\theta_1^{kM+1}+\theta_{\mathrm{a}}^{(k-1)M+1}) & \cdots & R_U^{kM+1}\sin(\theta_U^{kM+1}+\theta_{\mathrm{a}}^{(k-1)M+1}) \\ R_1^{kM+1}\cos(\theta_1^{kM+1}+\theta_{\mathrm{a}}^{(k-1)M+1}) & \cdots & R_U^{kM+1}\cos(\theta_U^{kM+1}+\theta_{\mathrm{a}}^{(k-1)M+1}) \end{bmatrix} \tag{6}$$

Although the subject may generate multiple target candidates, clutter may also exist in the TCM. Therefore, a range of values is specified based on the prior motion information of the target, which is very similar to the improved range bin selection method we proposed previously [24]. The distinction is that the specified range of values in this study is expanded from a one-dimensional line to a two-dimensional surface. All target candidates within a specific range are regarded as being produced by the subject and creating a new target matrix $\boldsymbol{A}_{\mathrm{TM}}^{kM+1}$. Then, averaging along each row of the $\boldsymbol{A}_{\mathrm{TM}}^{kM+1}$ yields a rotation decision matrix $\boldsymbol{A}_{\mathrm{DM}}^{kM+1}$ for that frame with one column and five rows, as shown in (7).

$$\boldsymbol{A}_{\mathrm{DM}}^{kM+1} = \begin{bmatrix} \overline{R}^{kM+1} \\ \overline{v}^{kM+1} \\ \overline{\theta}^{kM+1} \\ \overline{R}^{kM+1}\sin\left(\overline{\theta}^{kM+1}+\theta_{\mathrm{a}}^{(k-1)M+1}\right) \\ \overline{R}^{kM+1}\cos\left(\overline{\theta}^{kM+1}+\theta_{\mathrm{a}}^{(k-1)M+1}\right) \end{bmatrix}, \tag{7}$$

where $\overline{\theta}^{kM+1}$ is the angle between the subject and the positive direction of the radar module. To obtain the $\boldsymbol{A}_{\mathrm{TM}}^{kM+1}$, the specific range of values for the current TCM is defined as

1) The center coordinates of the circle are set to the values $\boldsymbol{A}_{\mathrm{DM}\ (4,1)}^{(k-1)M+1}$ and $\boldsymbol{A}_{\mathrm{DM}\ (5,1)}^{(k-1)M+1}$ of the preceding rotation period.

2) The human walking process can be approximated as a uniform motion, so the radius r_{s}^{kM+1} of the specific range of values can be calculated using (8) and (9). The protection unit β_{s} is also introduced to increase fault tolerance, considering that there is occasionally acceleration from the stationary state to the uniform state.

$$r_{\mathrm{s}}^{kM+1} = \left(\left|\overline{v}^{(k-1)M+1}\right| T_{\mathrm{r}}\right) + \beta_{\mathrm{s}}, \tag{8}$$

$$\beta_{\mathrm{s}} = \frac{1}{2} a_{\mathrm{s}} {T_{\mathrm{r}}}^2, \tag{9}$$

where a_{s} is the subject's acceleration, and the value can be adjusted based on the conditions of the environment and the application background.

3) Assume that matrix TCM has u ($u = 1, 2, 3, \ldots, U$) target candidates. Then, the target candidates satisfying (10) are retained to form $\boldsymbol{A}_{\mathrm{TM}}^{kM+1}$. This is compared to the matrix $\boldsymbol{A}_{\mathrm{TCM}}^{0}$ when k is zero.

$$r_{\mathrm{s}}^{kM+1} \geq \sqrt{\left(\boldsymbol{A}_{\mathrm{TCM}\ (4,u)}^{kM+1} - \boldsymbol{A}_{\mathrm{DM}\ (4,1)}^{(k-1)M+1}\right)^2 + \left(\boldsymbol{A}_{\mathrm{TCM}\ (5,u)}^{kM+1} - \boldsymbol{A}_{\mathrm{DM}\ (5,1)}^{(k-1)M+1}\right)^2} \tag{10}$$

When $\boldsymbol{A}_{\mathrm{TCM}}^{kM+1}$ or $\boldsymbol{A}_{\mathrm{TM}}^{kM+1}$ is empty, $\boldsymbol{A}_{\mathrm{DM}}^{kM+1}$ inherits the matrix $\boldsymbol{A}_{\mathrm{DM}}^{(k-1)M+1}$ except for the second element (instantaneous velocity information), because there is no moving target in a particular range. Hence, the value of the second element is zero.

By contrast, $\boldsymbol{A}_{\mathrm{TM}}^{kM+1}$ is averaged along the row direction to produce $\boldsymbol{A}_{\mathrm{DM}}^{kM+1}$ if it is not empty. $\boldsymbol{A}_{\mathrm{DM}\ (3,1)}^{kM+1}$ is the arrival angle of the subject, and its angle with the positive direction of the y-axis in absolute coordinates is $\boldsymbol{A}_{\mathrm{DM}\ (3,1)}^{kM+1} + \theta_{\mathrm{a}}^{(k-1)M+1}$.

The rotation angle threshold value θ_{s} is set to filter angle $\boldsymbol{A}_{\mathrm{DM}\ (3,1)}^{kM+1}$. If the subject's movement results in only a slight angle shift, it is not essential to rotate the motor to change the radar's orientation. Thus, the proposed method can be used to ensure that the subject is always within the measured range and not precisely on the midline of the radar azimuth. Frequent, brief back-and-forth rotations may reduce the accuracy of heart rate measurements. The Algorithm I calculates the angle $\theta_{\mathrm{r}}^{kM+1}$ the motor needs to rotate at each period and updates the angle $\theta_{\mathrm{a}}^{kM+1}$. Finally, the stepping motor operates according to angle $\theta_{\mathrm{r}}^{kM+1}$.

In the multi-subject scenario, regardless of whether the other subjects are stationary, their information will not be included in the $\boldsymbol{A}_{\mathrm{TM}}$ since they are not within a specific range. Usually, the specific range is very small. For example, if the rotation period is 0.2 s and the average speed is 1 m/s, the

Algorithm I Rotation Angle Calculation

θ_{step} is the step angle of the stepping motor (a constant parameter);

$\lceil \ldots \rceil$ rounds the element to the next larger integer ;

$\lfloor \ldots \rfloor$ rounds the element down to the next lower integer ;

If $A_{\text{DM}}^{kM+1}{}_{(3,1)} < \theta_{\text{s}}$ **then**

$\theta_{\text{r}}^{kM+1} \leftarrow 0$;

$\theta_{\text{a}}^{kM+1} \leftarrow \theta_{\text{a}}^{(k-1)M+1}$;

Else

If $A_{\text{DM}}^{kM+1}{}_{(3,1)} > 0$ **then**

$\theta_{\text{r}}^{kM+1} \leftarrow \left\lceil \frac{A_{\text{DM}}^{kM+1}{}_{(3,1)}}{\theta_{\text{step}}} \right\rceil \theta_{\text{step}}$;

Else

$\theta_{\text{r}}^{kM+1} \leftarrow \left\lfloor \frac{A_{\text{DM}}^{kM+1}{}_{(3,1)}}{\theta_{\text{step}}} \right\rfloor \theta_{\text{step}}$;

$\theta_{\text{a}}^{kM+1} \leftarrow \theta_{\text{a}}^{(k-1)M+1} + \theta_{\text{r}}^{kM+1}$;

Return: $\theta_{\text{r}}^{kM+1}, \theta_{\text{a}}^{kM+1}$;

radius of the search range is only about 0.2 m. The probability of other subjects entering this range is very low. An object may generate more than one target candidate point. Even if the other subjects are close to the tracked subject, only a few points from other subjects exist within the specific range. Then, the average processing in the calculation of $\boldsymbol{A}_{\text{DM}}$ will remove their effect. Therefore, the proposed tracking method can also correctly lock the target without interference from other subjects in a multi-subject scenario.

3.2 Heart Rate Monitoring

The approach mentioned above ensures that the subject is constantly within the measurement range. By examining the RPM, the phase change information $\varphi(t)$ can be obtained from the range bin cell where the subject is located. The thoracic skin displacement $x(t)$ is determined by using the phase change information, and this displacement is then utilized to extract the heartbeat signal and determine the heart rate.

As discussed in the introduction, subjects had to remain still (sitting, standing, or lying down) for the measurement in most earlier studies using radar-based heart rate monitoring. In addition, the subjects were typically close to the radar. In this situation, techniques such as RPM peak detection can quickly identify the range bin cell in which the subject is located. Furthermore, once the range bin cell is located, no further modifications are required for the duration of the measuring cycle.

In contrast, the subjects in this study were accompanied by two states, motion and stationary, implying that the target range bin cells were constantly changing. Additionally, because RPM peak-seeking or RDM peak-seeking is susceptible to noise, it is challenging to locate the target range bin cells using these methods alone. Therefore, we suggested the improved adaptive range bin selection approach [24], which is also utilized in this study, to collect high-quality phase information precisely and swiftly.

The initial location confirmation and adaptive range bin selection are two parts of the IARBS method. The initial DRM $\boldsymbol{D}_{\text{i}}$ is obtained in the first step by immediately accumulating the up-chirp signals of H frames, which ensures that the appropriate range bin cell $\alpha_{\text{optimal}}^{H}$ is acquired at the beginning.

In the second stage, starting with frame $H+1$, the instantaneous velocity of the subject, as determined by the data from the preceding frames, adaptively limits the peak-seeking range of the DRM for each frame. The following are the justifications for not directly using the data from matrix $\boldsymbol{A}_{\text{DM}}$. High real-time performance is needed for motor rotation, and the rotation angle calculation function quickly calculates the subject's approximate position per M frames. However, heart rate measurement must precisely obtain the optimal range bin cell.

$\text{a}_{\text{max}}^{H+j}$, $\alpha_{\text{min}}^{H+j}$, and $\alpha_{\text{optimal}}^{H+j}$ are the upper and lower bounds of the peak-seeking range and the outcomes for frame $H+j$ ($j = 1, 2, 3, \ldots, J$), respectively. Their relationship is defined by (11), (12), and (13).

$$\alpha_{\text{min}}^{H+j} = \alpha_{\text{optimal}}^{H+(j-1)} - \left\lceil \frac{v_{\text{m}}^{H+j} T_{\text{f}}}{R_{\text{bin}}} \right\rceil - \beta_{\text{p}}^{H+j}, \tag{11}$$

$$\alpha_{\text{max}}^{H+j} = \alpha_{\text{optimal}}^{H+(j-1)} + \left\lceil \frac{v_{\text{m}}^{H+j} T_{\text{f}}}{R_{\text{bin}}} \right\rceil + \beta_{\text{p}}^{H+j}, \tag{12}$$

$$\beta_{\text{p}}^{H+j} = \left\lceil \eta \left(\frac{\alpha_{\text{max}}^{H+j} - \alpha_{\text{min}}^{H+j} + 1}{2} \right) \right\rceil, \tag{13}$$

where $\lceil \ldots \rceil$ rounds the element to the next larger integer, and R_{bin} represents the length of each range bin cell. The instantaneous velocity of the subject from frame $H+(j-1)$ to frame $H+j$ is represented by v_{m}^{H+j}. The coefficient η ($0 \leq \eta \leq 1$) and the peak-seeking range determine the protection cell length, represented as β_{p}^{H+j}. H and η are typically set at 0.5 and 0.1, respectively, according to [24].

After choosing the optimal range bin cell, L phase information (L chirp signals per frame) can be collected for each frame and averaged to enhance the phase quality. This implies that each frame extracts a phase value in the optimal range bin to which it corresponds. Then, the phase information of each frame in an observation window is stitched together in order of time, which gives the phase information $\varphi(t)$ over time.

Sudden body shaking or noise may occur during the measurement, and longer observation windows will provide better immunity to interference. The measurement will also be more accurate because the longer the data, the higher is the resolution of the FFT bins. In contrast, a shorter observation window provides a better real-time performance. The length of the observation window should be adjusted according to the application. In addition, if the subsequent windows are consecutive, the first stage of the IARBS method can be skipped after processing for the first observation window. Finally, the subject's thoracic displacement $x(t)$ is determined using (3).

Along with the heartbeat signal, $x(t)$ includes the respiration signal and other noises (e.g., body and radar shaking). Wavelet transform and bandpass filter are two techniques for

extracting heartbeat signals [32]–[35]. However, these conventional techniques have limitations because the heartbeat signal differs from person to person and the human heart rate is close to the high-order harmonics of the respiratory frequency.

An improved empirical mode decomposition (EMD) method called ICEEMDAN can adaptively divide a signal into a limited number of intrinsic mode functions (IMFs) depending on the time scale of the signal [36], [37]. Therefore, this study uses the ICEEMDAN method to decompose the $x(t)$ and reconstruct the heartbeat signal, which was confirmed in our previous work. The relationship between $x(t)$ and each IMF, and the residual $r_n(t)$ is given by (14).

$$x(t) = \sum_{i=1}^{n} IMF_i(t) + r_n(t) \quad i = 1, 2, 3, \ldots, n \tag{14}$$

Then all IMFs perform a spectral analysis. The IMF with energy concentrated at 0.8 Hz to 2.0 Hz as heart IMF is extracted to reconstruct the heartbeat signal, and the heart rate can be calculated [28].

4. Experiment

4.1 Equipment

In this experiment, a stepping motor with the specification NEMA17 was employed, and an Arduino UNO development board was used to connect the motor to the computer. The angle information of the subject was transmitted from the computer to the Arduino via serial communication. This motor is a hybrid stepping motor that combines the benefits of reactive and permanent magnet types. It has a high resolution, speed, and torque, and the specific parameters are listed in Table 1.

The radar module used in this research was based on Texas Instruments Inc. IWR1443. It operates at frequencies in the 77–81 GHz range and can continuously chirp up to 4 GHz. However, the radar module can only send chirp signals with a bandwidth of 3.6 GHz during operation, owing to local laws in Japan. Furthermore, the maximum effective isotropic radiated power (EIRP) complies with the Japanese Radio Law and the FFC regulations. It does not cause harm to the human body. In contrast to the SISO mode used by traditional heart rate measurement techniques, in this study, the radar uses a 1Tx4Rx mode to gather data on the subject's angle. The main parameters of the radar module are listed in Table 2.

Although the MIMO mode provides a better angular resolution than the SIMO mode, the SIMO mode is utilized in this study for the following reasons.

First, the background of this application is the medical monitoring of older people living alone. In most cases, there is typically no interference from other moving targets after screening using range-FFT and Doppler-FFT processing.

Second, as discussed in Sect. 3, after locking the target subject using the proposed method, the range of the current target selection is restricted to the spatial plane based on the target position of the previous rotation period. Therefore, even if other subjects are present during the measurement procedure, the proposed method ensures that there is no interference from them.

Third, employing the 3Tx4Rx mode for radar will inevitably increase the frame period and the amount of computation. This will reduce the real-time performance of the motor rotation and lead to the possibility that the subject will be out of the measurement range. The SISO data is utilized to calculate the chest displacement; hence, switching to the MIMO mode while maintaining a constant frame time will reduce the length of the SISO data. The phase quality of the IF signal is subsequently reduced. To increase measurement accuracy, this study focus on the phase quality of the IF signal rather than just the angular precision and angular resolution of the target subject.

In addition, we proposed in [24] a measurement method to simultaneously measure the heart rate of multiple people while walking. The proposed method for this study tracks a target subject and measures the vital signs, aiming to maximize the measurement range. Theoretically, it can also obtain the vital signs of other people within the measurement range. Moreover, the motor rotation angle calculation algorithm can be initialized to change the target subject.

A physical diagram of the hardware is shown in Fig. 6. A bracket holds the radar module on a circular plate, and the stepping motor rotates the plate. A DC power supply feeds the motor.

Table 1 Motor parameters list.

Type	Bipolar stepping motor
Stepping angle	1.8 deg.
Stepping mode	Full-step
Hold torque	5500 g·cm
Rated voltage	12 V
Rated current	1.5 A

Table 2 Radar module main parameters.

Parameter	Value
Bandwidth	3.6 GHz
Sweep time	51.4 µs
Frame length	50 ms
Slope	70 MHz/µs

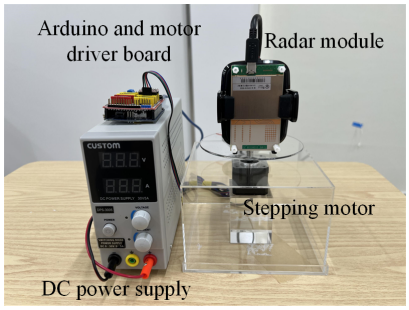


Fig. 6 The physical diagram of the hardware.

4.2 Experimental Environment and Route

The experiment for this study has four parts: a fixed route experiment, a random route experiment, an arm swing (more normal walking motion) comparison experiment, and a multi-subject scenario experiment.

There are five fixed routes, A, B, C, D, and E, as shown in Fig. 7(a) and (b). Routes A, B, and C are horizontal lines with midpoints of 0.5, 1, and 1.5 m from the radar module, and their ends are at an angle of 90 degrees from the radar. Move along a straight line with an average speed of roughly 0.5 m/s and 1 m/s, starting at the right endpoint of each route. For simplicity, the routes are denoted as A1, B1, C1 (average speed of 0.5 m/s), and A2, B2, and C2 (average speed of 1 m/s), respectively. The x-axis and y-axis directions are also defined in the various subplots of Fig. 7.

Route D is a 180-degree arc with a radius of one meter centered on the radar coordinates. The starting point coordinates are $(1, 0)$ and also move with an average speed of 0.5 m/s (route D1) and 1 m/s (route D2), respectively.

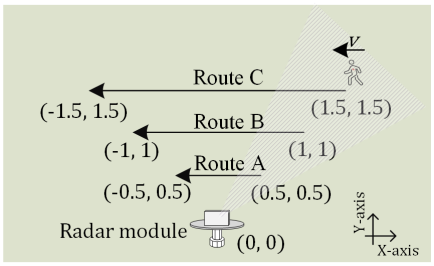
The above routes evaluate the feasibility of the proposed method, whereas route E is considered for its applicability. The subject traveled down route E in a 4 m $\times$ 2 m rectangle at a typical walking speed of 1 m/s. Furthermore, in the fixed route experiment, the subject attempted to face the radar module with either the front or back of the body.

The random route is then a free walk within a 4.5 $\times$ 5 square meter area in front of the radar module to further measure the practical application potential of the proposed method.

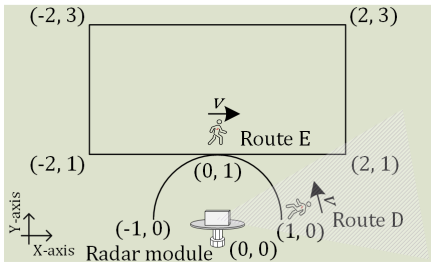
There are two routes for the arm swing comparison experiment, as shown in Fig. 7(d). Routes F1 and G1 indicate scenes without arm swings, whereas F2 and G2 indicate scenes with arm swings. The subjects moved along the route back and forth at an average speed of no more than 1 m/s.

The multi-subject scenario experiment aims to demonstrate that the proposed measurement system can precisely track the target subject in this case. This experiment was conducted with three subjects simultaneously, as shown in Fig. 7(e). Subject I, as the target, walked back and forth between 0.5 m and 3.5 m in front of the radar at an average speed of no more than 1 m/s (with the arm swing). Subjects II and III walked randomly along a horizontal straight line 1 and 3 m from the front of the radar, respectively. This experiment can observe whether the orientation of the radar is interfered with by other people while the measurement system tracks subject I.

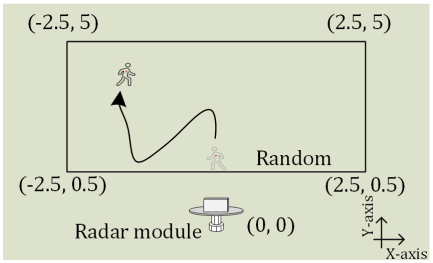
The subjects alternated between walking and constantly standing during each measurement period of 60 s to simulate the movement of people in the room. Figure 8 depicts the actual experimental scenario. Additionally, the subject wore an ECG device to record heart rate data during the measurement.



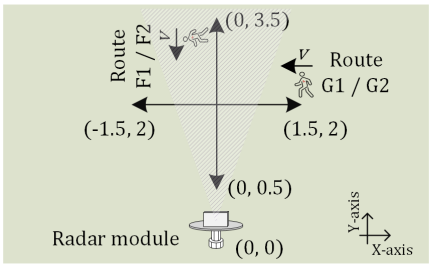
(a)



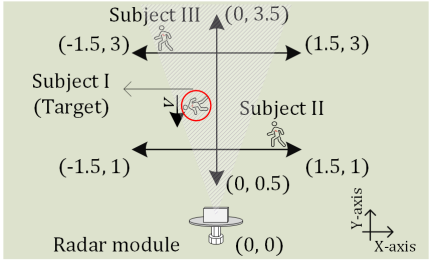
(b)



(c)



(d)



(e)

Fig. 7 Walking route maps of subjects.

(a) Non-swinging arm scenario

(b) Arm swing scenario

(c) Multi-subject scenario

Fig. 8 Experiment scenarios.

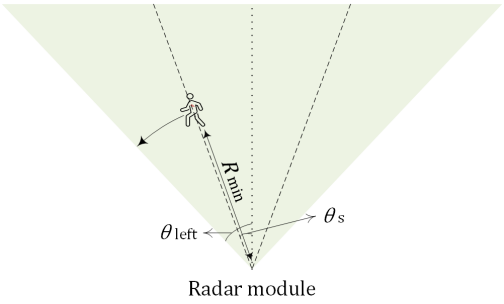


Fig. 9 The figure depicts how to select the appropriate minimum rotation angle by the half-power angle of radar.

4.3 Parameter Analysis

The radar's horizontal left-half power angle is θ_{left}, and the minimum measurement distance is $R_{\min}$. The following equation must be satisfied by the variables M and θ_s.

$$MT_f v_p < \frac{|\theta_{\text{left}} - \theta_s|\,\pi R_{\min}}{180^\circ}, \tag{15}$$

where v_p is the speed at which the person moves under normal conditions. As shown in Fig. 9, assume that the subject's distance from the radar is precisely $R_{\min}$ and that the subject's angle with the radar centerline is infinitesimally close to θ_s. In the next rotation period, the possible walking distance of the subject is at most $MT_f v_p$ to ensure that it is within the measurement range. In the context of the application of this measurement system, elderly individuals who live alone typically do not move indoors at speeds greater than 1 m/s. This experiment sets M and θ_s to 4 and 5 degrees because $R_{\min}$ is 0.5 m and the half-power angle of azimuth and elevation are about 70 and 10 degrees, respectively.

Table 3 Information about subjects.

Subject	Height [cm]	Weight [kg]
A	175	60
B	170	56
C	169	60
D	174	61
E	176	62
F	166	78

4.4 Experimental Results

Six volunteers participated in this experiment, and Table 3 lists their heights and weights. In order to better evaluate the proposed approach, it is necessary to reduce the variable factors between the different subjects as much as possible. Therefore, the height of the radar module was adjusted according to the thoracic height of the volunteer before the beginning of the measurement to try to keep both at the same horizontal plane. Each volunteer participated in all the routes of this experiment.

Figure 10 shows the cumulative distribution functions (CDF) of the absolute values of the measured AoA ($\left|\text{Angle}\right| : \left|\boldsymbol{A}_{\text{DM}}^{kM+1}{}_{(3,1)}\right|$) for each rotation period of the fixed route and the random route experiments for all subjects. Overall, the subjects remained within the ideal measurement range of plus or minus 35 degrees in front of the radar throughout the measurement cycle.

Owing to the subjects' faster angular velocity, as they moved along route A, the AoA values for route A were greater than those measured for the other routes. The results for routes B and C showed that in 90 percent of cases, the subjects had an AoA of less than 10 degrees. The AoA values increased with the movement speed of the subject. However, the effect is not particularly substantial, according to the combined results of all fixed route experiments.

The results for route D show that the measurement system tracks accurately even when the subject's angle changes by up to 180 degrees. Accordingly, depending on the requirements of the application, it is possible to position the radar module in the center of the room for 360-degree tracking measures. The angular velocity of the subject relative to the radar during route D was higher than that of routes A, B, and C, which led to a more significant difference in the results of the CDF of routes D1 and D2 than the other routes. In addition, the measurement system can accurately track the subject during random walks. In a few cases, the angle of the subject's position was approximately 30 degrees. However, no subject's position was outside the ideal measurement range, which demonstrates the reliability of the system.

Because route E contains straight lines similar to routes A-C, Fig. 11 shows only the results of the traced trail for one

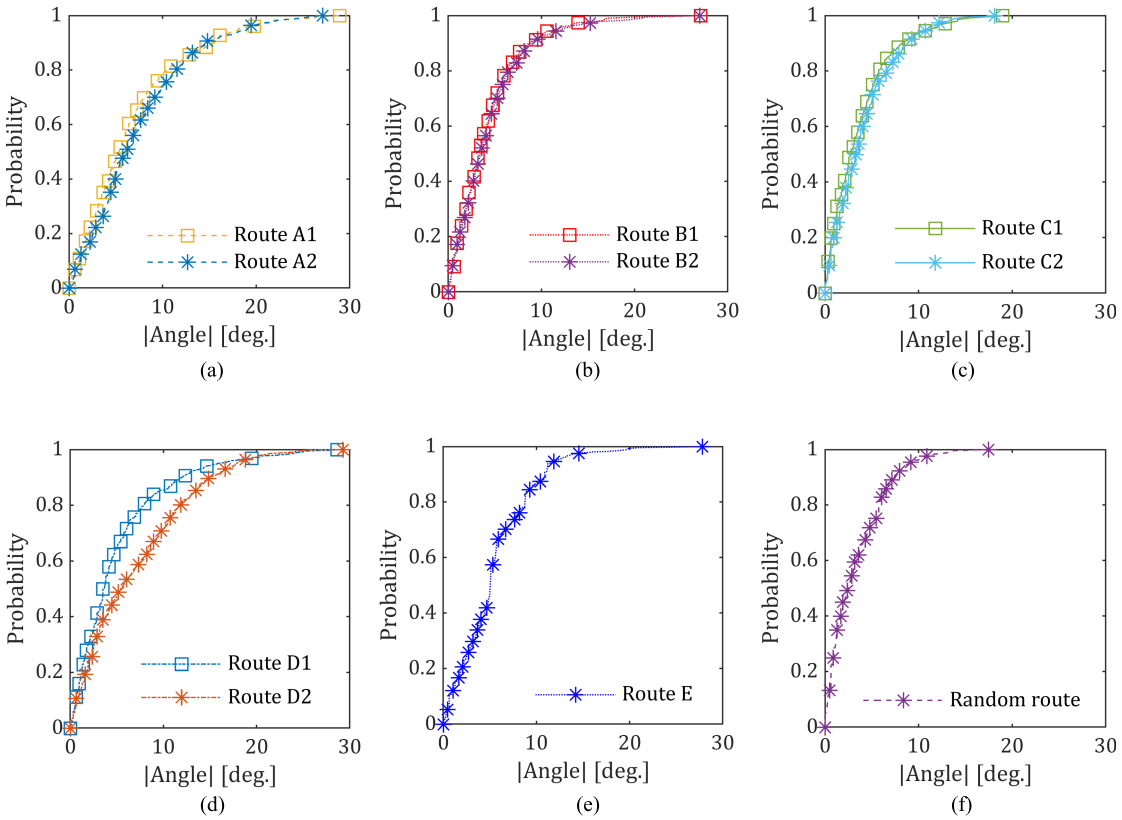


Fig. 10 The cumulative distribution functions of the absolute values of the measured subject's angle information.

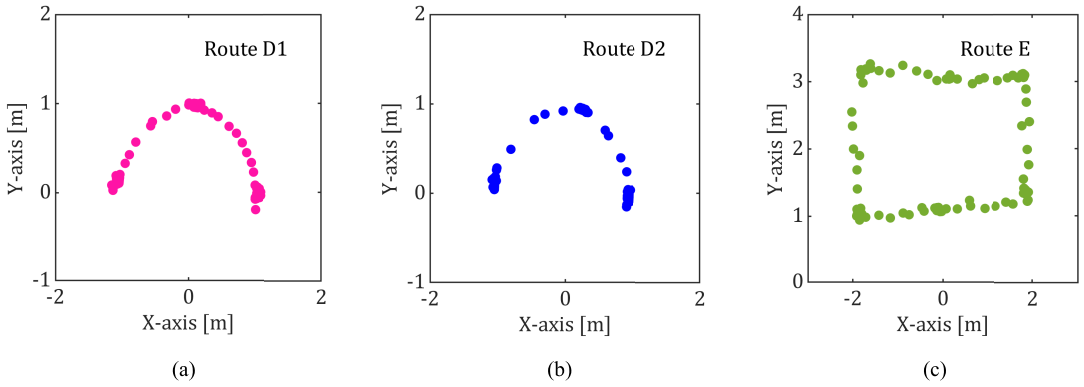


Fig. 11 The results of the traced trail for one subject walking along routes D and E.

subject walking along routes D and E. The results show that the tracking trail is still rather precise, even if the proposed method quickly detects the subject's approximate location and sends it to the motor.

The RPM of a subject walking along route D2 and the optimal range bin results achieved by the IARBS method are shown in Fig. 12(a) and (b). The phases extracted from the optimal range bins are then stitched together along the slow time dimension to calculate the $x(t)$ of the subject, as shown in Fig. 12(c).

Next, $x(t)$ was decomposed into a finite number of IMFs using the ICEEMDAN method. FFT processing is performed on them to select the heart IMF among IMFs. The analysis was performed with a 30 s observation window. The results after decomposition by ICEEMDAN are shown in Fig. 13, with the time and frequency domains of each IMF on the left and right, respectively. According to the frequency domain analysis, the heart IMF and breathing IMF were IMF_3 and IMF_5. This study focused on the heartbeat signal because the breathing signal is a low-frequency component that is simple to extract. Finally, the heartbeat signal was reconstructed. Finally, IMF_3 was used to reconstruct the heartbeat signal and estimate the heart rate.

In addition, the phase change data and low-frequency

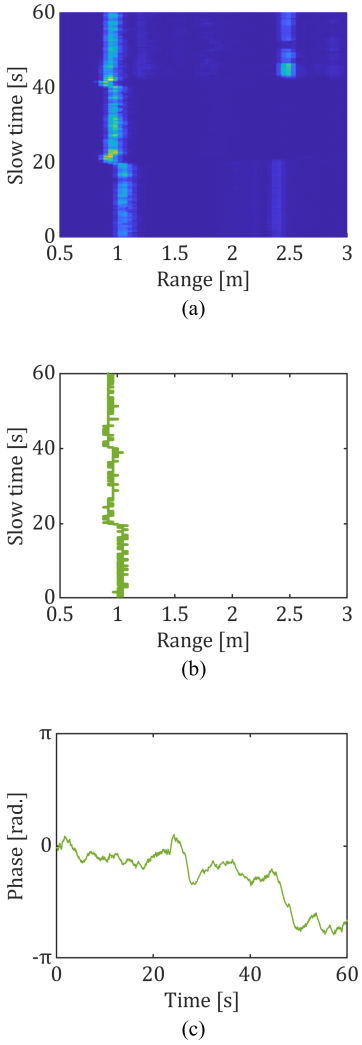


Fig. 12 The RPM (a) of a subject walking along route D2, the optimal range bin results (b) achieved by the IARBS method, and the stitched phase change information (c) are shown in this figure.

IMFs demonstrate that walking causes the $x(t)$ signal to vary more than when the subject is motionless [28].

Figure 14(a) shows a frequency domain comparison of the reconstructed heartbeat signal and ECG signal. These measurement data were obtained using route D2. The heart rates obtained by the proposed method and ECG were 1.767 Hz and 1.833 Hz, respectively, with an absolute error of approximately 3.96 beats per minute (bpm), and the accuracy was approximately 96.4%.

The spectrum of the measurements is wider than that of the ECG. The following are some of the possible reasons for this phenomenon. Firstly, there is a fundamental difference between ECG, an electrical signal, and radar-based heart rate measurement, measured by a displacement signal from the skin. Secondly, the human chest skin is not plane, and the skin displacement caused by the heartbeat is slightly different at each location. Thirdly, the subject was walking, and there was inevitably a small amount of noise compared to stationary. The ICEEMDAN method has minimized the

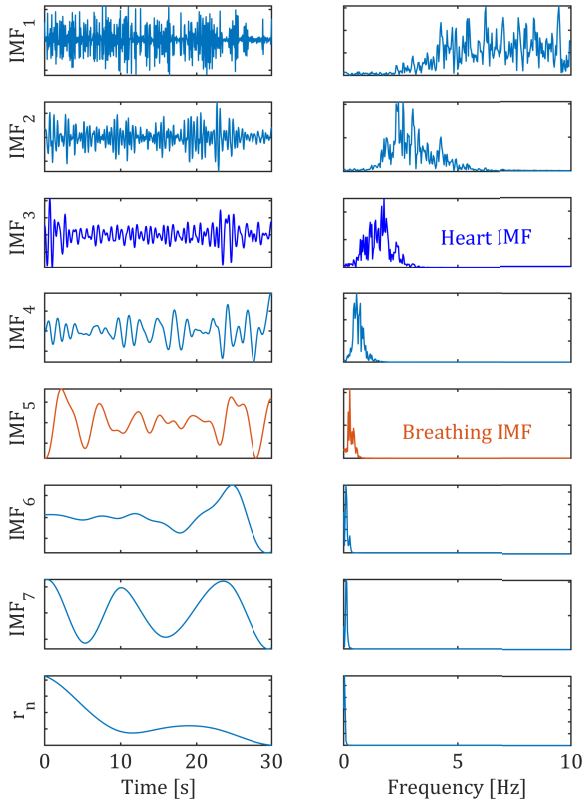


Fig. 13 Results of ICEEMDAN decomposition. The left and right sides are the time domain and frequency domain, respectively.

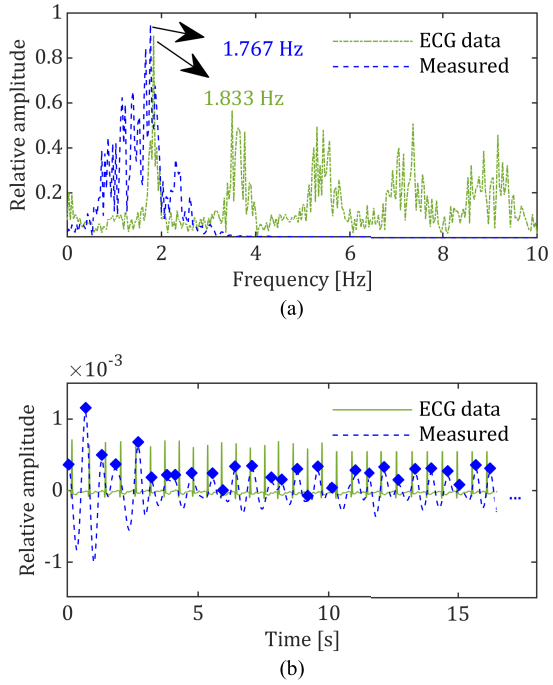


Fig. 14 (a) The frequency domain comparison of the reconstructed heartbeat signal and the ECG signal. (b) The time domain comparison of the reconstructed heartbeat waveforms and the ECG waveform.

Table 4 Accuracy of heart rate measurement [%].

Route	A1	A2	B1	B2	E
Mean	98.42	97.47	97.29	96.78	97.19
Route	C1	C2	D1	D2	Random
Mean	97.69	96.64	98.06	97.08	96.16
Route	F1	F2	G1	G2	Multi-subject
Mean	97.05	96.75	97.83	97.27	96.22

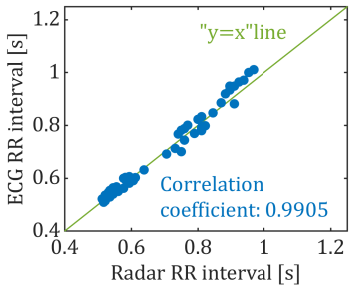


Fig. 15 The correlation of the RR interval measured by radar with ECG data.

amplitude of the interference term as much as possible. In this research field, the most attention is paid to the accuracy of heart rate and RR interval measurements. Therefore, the practicality of the proposed method is not affected.

Table 4 lists the average heart rate measurement accuracy for each route. The accuracy increased with proximity to the radar or slower speeds, and route A1 had the highest measurement accuracy at 98.42%. The results of routes F and G show that swinging the arm while walking decreases the accuracy of the heart rate measurement by approximately 0.3% to 0.56%. The accuracy of the heart rate measurements decreased by 0.53%, while other people interfered. The overall average measurement accuracy, including the random route, was above 96%, thus maintaining a high level of accuracy.

To better observe the details, a portion of the reconstructed heartbeat waveform is compared with its corresponding ECG waveform in the time domain, as shown in Fig. 14(b). The results showed a higher correlation between the measured heartbeat waveform's RR interval (spacing between rhombuses) and the ECG data. Meanwhile, Fig. 15 depicts the correlation between the mean RR interval time of all measurements of the fixed route experiment, random route experiment, and the corresponding ECG. The radar and ECG data were the horizontal and vertical coordinates, respectively. The correlation coefficient γ of the two datasets was calculated using (16), and their correlation coefficient was 0.9905, indicating a high correlation between them.

$$\gamma = \frac{1}{P-1}\sum_{i=1}^{P}\left(\frac{F_i-\mu F}{\sigma F}\right)\left(\frac{E_i-\mu E}{\sigma E}\right), \tag{16}$$

where P and i are the total number of data and their indices, respectively. F is the data measured by the radar and E is the ECG data used as a reference. The means of the two data are μF and μE and their standard deviations are σF and σE. Figure 16 shows the CDF of the absolute error of all fixed

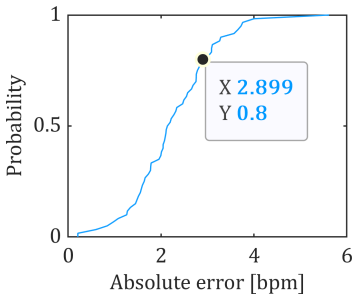


Fig. 16 The cumulative distribution functions of the absolute error of heart rate measurement.

Table 5 The RMSE of heart rate measurement (the fixed and random route experiment) [BPM].

Route \ Subject	A	B	C	D	E	F	Mean
A1	1.99	1.75	1.88	2.05	1.89	1.79	1.89
A2	3.71	3.38	3.78	2.68	3.94	1.86	3.23
B1	2.57	2.55	3.66	3.58	3.52	1.14	2.84
B2	3.99	3.82	4.67	3.85	3.25	3.74	3.89
C1	2.30	2.34	3.67	3.32	2.00	2.75	2.73
C2	5.32	3.03	4.25	3.72	3.86	4.03	4.04
D1	2.46	2.77	3.02	1.78	2.67	2.01	2.45
D2	2.64	3.91	3.83	5.23	3.05	3.44	3.68
E	4.01	2.42	3.73	5.39	2.45	4.00	3.67
Random	6.64	2.31	6.82	6.33	2.58	1.42	4.35

and random route experiment measurements. In 80% of the measurements, the absolute error was under 2.9 bpm. There were no instances when the absolute error exceeded 6 bpm, even when the random route results were considered.

To further quantitatively analyze the proposed method's heart rate measurement results, 60 s streaming data were analyzed with a 30 s observation window and a 1 s sliding step. Then, the heart rate for each observation window is calculated and contrasted with the ECG data to determine the root mean square error (RMSE), as shown in (17).

$$\mathrm{RMSE} = \sqrt{\frac{1}{Q}\sum_{i=1}^{Q}(HRR_i - HRE_i)^2}, \tag{17}$$

where Q and i are the total number of observation windows and their index numbers, respectively. *HRR* and *HRE* are the heart rates measured by each observation window and reference heart rate (ECG data), respectively.

Table 5 presents the RMSE of the heart rate measurements of the fixed and random route experiments. The results demonstrate that as the distance and speed increase, the RMSE of the heart rate estimation also increases. According to the average RMSE value for each route, the RMSE is approximately 1.2 bpm lower when the subject walks at 0.5 m/s as opposed to 1 m/s. We speculate that precision may be affected because the stepping motor rotates more frequently, causing the fixed plate to vibrate and generate noise. Meanwhile, the experimental results of [24] demonstrated that the speed of the subject can affect how accurately the IARBS method chooses the optimal range bin. Additionally, the positive association between RMSE and

Table 6 The RMSE of heart rate measurement (the arm swing comparison experiment) [BPM].

Route \ Subject	A	B	C	D	E	F	Mean
F1 (No arm swinging)	3.46	1.64	2.59	3.13	3.49	3.59	2.98
F2 (Arm swinging)	3.67	3.45	2.50	3.19	3.85	4.18	3.47
G1 (No arm swinging)	3.44	3.26	4.02	3.41	3.95	4.00	3.68
G2 (Arm swinging)	4.37	3.32	4.75	2.77	4.06	4.32	3.93

distance will rise less quickly or stop being significant if the subject's distance from the radar exceeds 1 m. Likewise, when the subject moves at a speed of 1 m/s, the rate at which the RMSE increases with distance will decrease. Therefore, the experimental results showed that the RMSE values of the proposed method for heart rate estimation converged to a particular value and were stable.

The heart rate estimates are highly accurate even when the subject's angle information changes significantly, as evidenced by the mean RMSEs of 2.45 bpm and 3.68 bpm for routes D1 and D2, respectively. The results for each volunteer in the random route experiment varied significantly. However, the RMSEs were all below 6.9 bpm, with a mean value of 4.35 bpm. Some of the larger RMSEs are caused by spending extended periods with the side of the body facing the radar, which decreases the measurement accuracy.

Table 6 presents the RMSE of the arm swing comparison experiment's heart rate measurements. Notwithstanding individual variations, swinging of the subject's arm throughout the measurement process can affect the measurement accuracy. During the arm swing, the average RMSE of routes F and G decreased by 0.49 bpm and 0.25 bpm, respectively. This influence is caused by the left anterior thoracic muscle movement when the arm is swung, and the movement of the upper arm also causes a change in the phase of the IF signal.

Figure 17(a), (b), and (c) depict the RPM of the multi-subject scenario experiment, the RDM of a certain frame during the measurement, and the optimal range bin results of the target subject achieved by the IARBS method, respectively. The target subject was within the measurement range at all times and the rotation of the radar module was not disturbed by other people. Figure 17(b) shows the RDM when the target subject and other subjects overlapped on the RPM. The IARBS method can adaptively adjust the peak-seeking range based on the velocity of the target subject in the previous time. It is robust to other subjects suddenly entering its peak-seeking range. Even if other subjects are at the same range bin as the target subject for an extended time, their heart rates can be separated by the ICEEMDAN method. Figure 17(c) shows that the IARBS method can accurately select the optimal range bin of the target subject in the multi-subject scenario. However, there are many interference factors in this case, such as the possible presence of multiple subjects' arm swings and body shaking simultaneously in one range bin cell.

As discussed in Sect. 3, longer observation windows result in better interference immunity. Figure 18 compares the heart IMF and ECG data in the frequency domain for each observation window in the multi-subject scenario experiment. The ECG data showed that the mean heart rate during the measurement period was 1.402 Hz. No significant peaks in the heart IMF made it challenging to estimate the heart rate, whether the observation window was 30, 40, or 50 s. The peak frequency is 1.417 Hz when the observation window is 60 s, and the absolute error is about 0.9 bpm compared to the ECG data, with an accuracy of 98.93%. The measurement accuracies of the six subjects were 95.37%, 98.93%, 95.55%, 96.28%, 95.49%, and 95.69%, respectively, with a mean value of 96.22%. The measurement accuracy of this experiment was also influenced by the random movements of the non-target subjects, which were similar to the random route experiment. The experimental results show that the heart rate measurement for the target subject

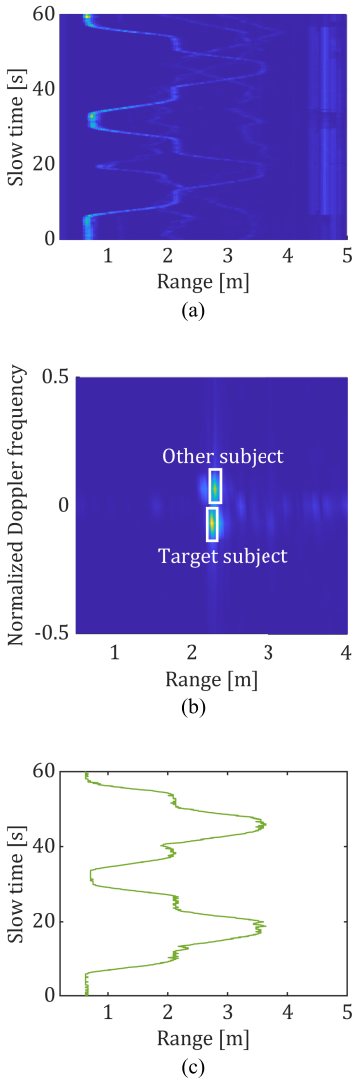


Fig. 17 The RPM (a) of the multi-subject scenario experiment, the RDM (b) of a certain frame during the measurement, and the optimal range bin results (c) of the target subject achieved by the IARBS method are shown in this figure.

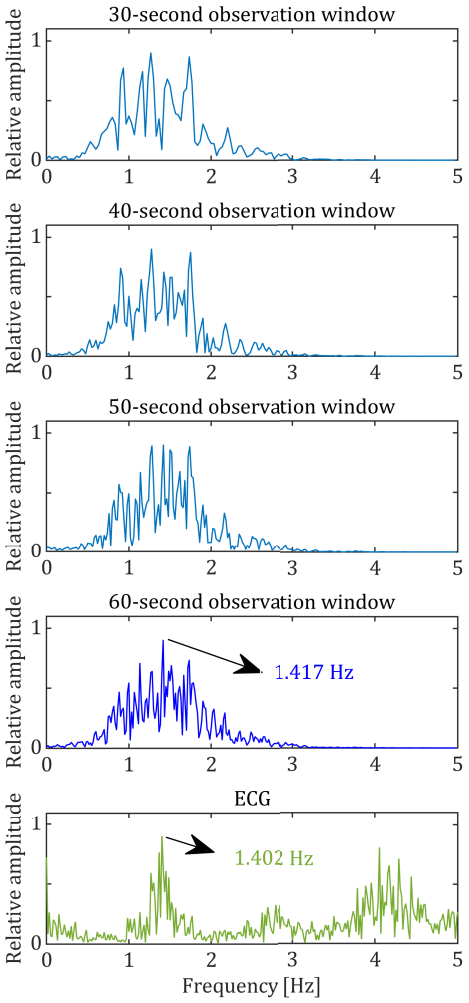


Fig. 18 Comparison of heart IMF and ECG data in the frequency domain for each observation window in the multi-subject scenario experiment.

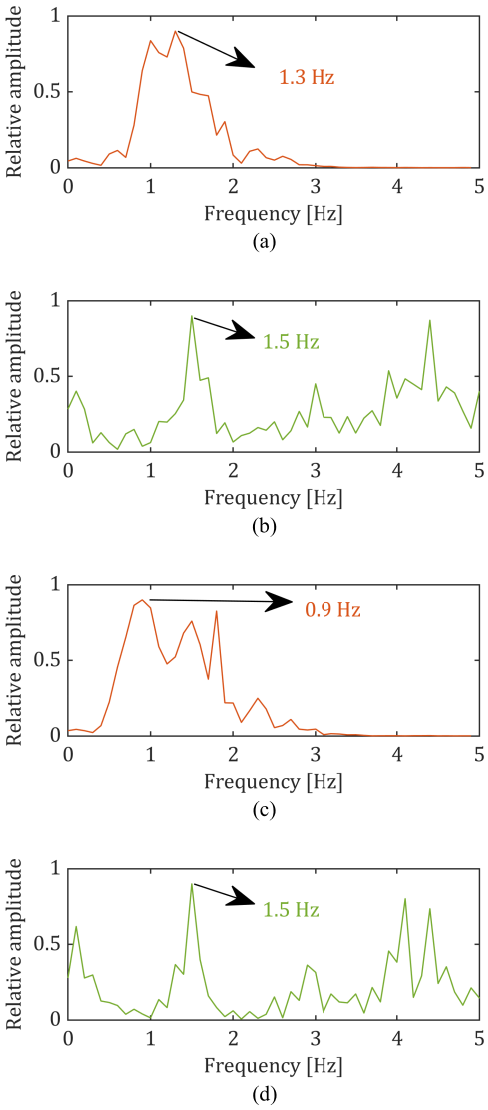


Fig. 19 In (a) and (c), the two spectrums of heart IMF were measured with a 10-second observation window, and the ECG data corresponding to them are shown in (b) and (d).

still maintains a high accuracy in multiple people scenarios.

In addition, it is challenging to improve real-time performance by shortening the observation window whenever heart rate is measured using FFT-based time-frequency domain analysis, including the ICEEMDAN method. For example, when the frame period is 0.05 s (typical), and the observation window is below 10 s and 5 s, the spectrum's frequency resolution will be as low as 0.1 Hz and 0.2 Hz. Meanwhile, the EMD-based algorithm (ICEEMDAN) is required for the data's length, which is not sufficient to separate the individual frequency components if it is too short. [15], [28] using EMD-based algorithms to measure heart rate with the subject at stationary have employed an observation window of at least 15 s. Since the subject in this study was moving, the evaluation was attempted with an observation window of 30 s or more.

Figure 19 shows two spectrums of heart IMF measured with a 10-second observation window. As shown in Fig. 19(a) and (b), the heart rate can be easily detected when interference is low in the spectrum of the heart IMF. However, the low frequency resolution of the spectrum resulted in a measurement error of 0.2 Hz. Due to the shortness of the data, heart rate estimation will become difficult when ICEEMDAN cannot efficiently extract high-quality, low-interference heart IMF, as shown in Fig. 19(c) and (d). Interference occurs randomly, and clearly, longer observation windows possess better robustness.

In order to trade off the real-time performance and stability of the measurement system, the following attempts are considered in future work. One is to shorten the frame period as much as possible to increase the data length, provided the radar hardware supports it. The other one is that the application uses an overall delayed dynamic display when displaying heartbeat waveforms. The next measurement and processing are performed simultaneously during the previous waveform display. When the previous waveform is just about finished, the next waveform is displayed immediately afterward.

Because this study is the first to estimate the heart rate by adjusting the radar measurement range through stepping motors, the measurement results are difficult to compare quantitatively with other papers. Therefore, the main focus is the qualitative comparison and analysis of this experiment's results with those of other works' measurements.

Most radar-based heart rate estimation methods, such as those in [2]–[20], cannot perform tracking measurements on moving targets, whereas the proposed method offers this possibility.

In the experiment in [23], the subject moved slowly and the heart rate was measured with 90% accuracy. In [21], the subjects made only backward and forward movements with a measurement error range of 0–8 bpm. By contrast, the subjects in this experiment were faster, and the RMSEs of the measurements were all below 7 bpm. Moreover, the average measurement accuracy for each route exceeded 96%. The measurement accuracy in our previous study [24] was 95.88%. This is because the subject was consistently inside the ideal measurement range in this experiment, which increased the accuracy.

The six participants in [15] were in a stationary state. At distances of 1 m, 1.5 m, 2 m, and 2.5 m from the radar, the mean values of the RMSE for heart rate estimation were approximately 2.39 bpm, 2.57 bpm, 3.23 bpm, and 4.69 bpm, respectively. Although the results for routes B and C were higher than 2.39 and 2.57 bpm, the farthest distance in route E was about 3.61 m, and the mean value of RMSE was only 3.67 bpm. In addition, the proposed method can change the radar orientation to adjust the measurement range and estimate the heart rate of a moving person, making it more advantageous for application.

5. Conclusion

In conclusion, this study proposes a novel radar-based adaptive tracking method for measuring the heart rate of a moving subject. The proposed algorithm is employed to determine the position of the subject to control a stepping motor that adjusts the radar measurement range. The results of the fixed-route experiments revealed that when the subject was moving at a speed of 0.5 m/s, the mean values of RMSE for heart rate measurements were all below 2.85 bpm, and when moving at a speed of 1 m/s, they were all below 4.05 bpm. When subjects walked at random routes and speeds, the RMSE of the measurements were all below 6.85 bpm, with a mean value of 4.35 bpm. In addition, this study not only evaluated the potential effect of arm swing (more normal walking motion) on heart rate measurement but also demonstrated the ability of the proposed method to measure heart rate in a multiple-person scenario.

Meanwhile, the overall measurement accuracy was greater than 96%, when the random route was included. Moreover, the RR intervals of the reconstructed heartbeat signal and ECG data were highly correlated, with a correlation coefficient of 0.9905. In the future, we plan to improve the experimental platform to reduce the noise generated by the radar as it rotates to improve the accuracy of the measurements.

Acknowledgments

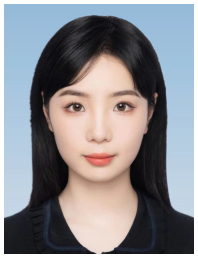
This study was supported by the School of Science and Technology, College of Science and Technology, Nihon University. This study involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was granted by the Nihon University Research Ethics Guideline. The authors would like to thank Kosuke Otsu for his help in collecting the experimental data.

References

[1] S. Yasuda, Y. Miyamoto, and H. Ogawa, "Current status of cardiovascular medicine in the aging society of Japan," Circulation, vol.138, no.10, pp.965–967, Sept. 2018.
[2] M. Alizadeh, G. Shaker, J.C.M.D. Almeida, P.P. Morita, and S. Safavi-Naeini, "Remote monitoring of human vital signs using mm-wave FMCW radar," IEEE Access, vol.7, pp.54958–54968, 2019.
[3] Y. Xiong, Z. Peng, C. Gu, S. Li, D. Wang, and W. Zhang, "Differential enhancement method for robust and accurate heart rate monitoring via microwave vital sign sensing," IEEE Trans. Instrum. Meas., vol.69, no.9, pp.7108–7118, Sept. 2020.
[4] L. Qiao X. Li, B. Xiao, M. He, X. Bi, W. Li, and X. Gao, "Learning-refined integral null space pursuit algorithm for noncontact multi-subjects vital signs measurements using SFCW-UWB and IR-UWB radar," IEEE Trans. Instrum. Meas., vol.71, pp.1–13, 2022.
[5] M. Mercuri, Y. Liu, I. Lorato, T. Torfs, F. Wieringa, A. Bourdoux, and C.V. Hoof, "A direct phase-tracking Doppler radar using wavelet independent component analysis for non-contact respiratory and heart rate monitoring," IEEE Trans. Biomed. Circuits Syst., vol.12, no.3, pp.632–643, June 2018.
[6] D. Yang, Z. Zhu, and B. Liang, "Vital sign signal extraction method based on permutation entropy and EEMD algorithm for ultra-wideband radar," IEEE Access, vol.7, pp.178879–178890, Dec. 2019
[7] X. Shang, J. Liu, and J. Li, "Multiple object localization and vital sign monitoring using IR-UWB MIMO radar," IEEE Trans. Aerosp. Electron. Syst., vol.56, no.6, pp.4437–4450, Dec. 2020.
[8] H. Shen, C. Xu, Y. Yang, L. Sun, Z. Cai, L. Bai, E. Clancy, and X. Huang, "Respiration and heartbeat rates measurement based on autocorrelation using IR-UWB radar," IEEE Trans. Circuits Syst. II, Exp. Briefs, vol.65, no.10, pp.1470–1474, Oct. 2018.
[9] P. Wang, F. Qi, M. Liu, F. Liang, H. Xue, Y. Zhang, H. Lv, and J. Wang "Noncontact heart rate measurement based on an improved convolutional sparse coding method using IR-UWB radar," IEEE Access, vol.7, pp.158492–158502, Oct. 2019.
[10] W. Xia, Y. Li, and S. Dong, "Radar-based high-accuracy cardiac activity sensing," IEEE Trans. Instrum. Meas., vol.70, pp.1–13, 2021.
[11] V.L. Petrović, M.M. Janković, A.V. Lupšić, V.R. Mihajlović and J.S. Popović-Božović, "High-accuracy real-time monitoring of heart rate variability using 24 GHz continuous-wave Doppler radar," IEEE Access, vol.7, pp.74721–74733, 2019.
[12] J. Park, Y. Jeong, G. Lee, J. Oh, and J. Yang, "915-MHz continuous-wave Doppler radar sensor for detection of vital signs," Electronics, vol.8, no.5, 561, 2019.
[13] M. Mercuri, Y. Liu, I. Lorato, T. Torfs, F. Wieringa, A. Bourdoux, and C.V. Hoof, "A direct phase-tracking Doppler radar using wavelet independent component analysis for non-contact respiratory and heart rate monitoring," IEEE Trans. Biomed. Circuits Syst., vol.12, no.3, pp.632–643, June 2018.
[14] N.T. Phuong Nguyen, P.-Y. Lyu, M.H. Lin, C.-C. Chang, and S.-F.

Chang, "A short-time autocorrelation method for noncontact detection of heart rate variability using CW Doppler radar," 2019 IEEE MTT-S International Microwave Biomedical Conference (IMBioC), pp.1–4, 2019.
[15] L. Sun, S. Huang, Y. Li, C. Gu, H. Pan, H. Hong, and X. Zhu, "Remote measurement of human vital signs based on joint-range adaptive EEMD," IEEE Access, vol.8, pp.68514–68524, April 2020.
[16] X. Zhang, Z. Liu, Y. Kong, and C. Li, "Mutual interference suppression using signal separation and adaptive mode decomposition in noncontact vital sign measurements," IEEE Trans. Instrum. Meas., vol.71, no.4001015, pp.1–15, 2022.
[17] T.K.V. Dai, K. Oleksak, T. Kvelashvili, F. Foroughian, C. Bauder, P. Theilmann, A.E. Fathy, and O. Kilic, "Enhancement of remote vital sign monitoring detection accuracy using multiple-input multiple-output 77 GHz FMCW radar," IEEE J. Electromagn. RF Microw. Med. Biol., vol.6, no.1, pp.111–122, March 2022.
[18] J. Liu, Y. Li, C. Li, C. Gu, and J.-F. Mao, "Accurate measurement of human vital signs with linear FMCW radars under proximity stationary clutters," IEEE Trans. Biomed. Circuits Syst., vol.15, no.6, pp.1393–1404, Dec. 2021.
[19] G. Wang, J. Muñoz-Ferreras, C. Gu, C. Li and R. Gómez-García, "Application of linear-frequency-modulated continuous-wave (LFMCW) radars for tracking of vital signs," IEEE Trans. Microw. Theory Techn., vol.62, no.6, pp.1387–1399, June 2014.
[20] G. Wang, C. Gu, T. Inoue, and C. Li, "A hybrid FMCW-interferometry radar for indoor precise positioning and versatile life activity monitoring," IEEE Trans. Microw. Theory Techn., vol.62, no.11, pp.2812–2822, Nov. 2014.
[21] Y. Rong, K.V. Mishra, and D.W. Bliss, "Multiple moving targets heartbeat estimation and recovery using multi-frequency radars," 2021 IEEE Radar Conference (RadarConf21), pp.1–5, 2021.
[22] X. Yang, X. Zhang, H. Qian, Y. Ding, and L. Zhang, "MMT-HEAR: Multiple moving targets heartbeats estimation and recovery using IR-UWB radars," 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp.5733–5736, 2020.
[23] A. Morimatsu, S. Matsuguma, and A. Kajiwara, "Heart rate estimation of a moving person using 79 GHz-band UWB radar," 2019 IEEE Sensors Applications Symposium (SAS), Sophia Antipolis, France, pp.1–5, May 2019.
[24] Y. Hu and T. Toda, "Remote vital signs measurement of indoor walking persons using mm-wave FMCW radar," IEEE Access, vol.10, pp.78219–78230, July 2022.
[25] P. Zhao, C.X. Lu, B. Wang, C. Chen, L. Xie, M. Wang, N. Trigoni, and A. Markham, "Heart rate sensing with a robot mounted mmWave radar," 2020 IEEE International Conference on Robotics and Automation (ICRA), pp.2812–2818, Aug. 2020.
[26] K. Gupta, Srinivas M.B., Soumya J, O.J. Pandey, and L.R. Cenkeramaddi, "Automatic contact-less monitoring of breathing rate and heart rate utilizing the fusion of mmWave radar and camera steering system," IEEE Sensors J., vol.22, no.22, pp.22179–22191, Nov. 2022.
[27] X. Peng, Y. Hu, and T. Toda, "Design and evaluation of remote tracking heart rate measurement based on millimeter-wave FMCW radar," IEICE Commun. Express, vol.11, no.6, pp.330–335, June 2022.
[28] Y. Hu and T. Toda, "The effect of multi-directional on remote heart rate measurement using PA-LI joint ICEEMDAN method with mm-wave FMCW radar," IEICE Trans. Commun., vol.E105-B, no.2, pp.159–167, Feb. 2022.
[29] Q. Zheng, L. Yang, J. Li, T. Hu, J. Zhu, C. Song, and Z. Xu, "A target detection scheme with decreased complexity and enhanced performance for range-Doppler FMCW radar," IEEE Trans. Instrum. Meas., vol.70, no.8001113, pp.1–13, 2021.
[30] V.V. Tien, T. Vu Hop, L.H. Nam, N. Van Loi, and T.T. Thanh, "An adaptive 2D-OS-CFAR thresholding in clutter environments: Test with real data," 2018 5th International Conference on Signal Processing and Integrated Networks (SPIN), pp.116–119, Sept. 2018.
[31] M. Kronauge and H. Rohling, "Fast two-dimensional CFAR procedure," IEEE Trans. Aerosp. Electron. Syst., vol.49, no.3, pp.1817–1823, July 2013.
[32] D.R. Morgan and M.G. Zierdt, "Novel signal processing techniques for Doppler radar cardiopulmonary sensing," Signal Processing, vol.89, no.1, pp.45–66, Jan. 2009.
[33] T.K.V. Dai, Y. Yu, P. Theilmann, A.E. Fathy, and O. Kilic, "Remote vital sign monitoring with reduced random body swaying motion using heartbeat template and wavelet transform based on constellation diagrams," IEEE J. Electromagn. RF Microw. Med. Biol., vol.6, no.3, pp.429–436, Jan. 2022.
[34] S. Sato, Y. Hu, and T. Toda, "A study of multi-directional heart-rate-estimation with discrete wavelet transform and band pass filter with 77 GHz-band FMCW radar," IEICE 1st Int. Conf. on Emerging Technologies for Communs. (ICETC), G1-4, Dec. 2020.
[35] M. Li and J. Lin, "Wavelet-transform-based data-length-variation technique for fast heart rate detection using 5.8-GHz CW Doppler radar," IEEE Trans. Microw. Theory Techn., vol.66, no.1, pp.568–576, Jan. 2018.
[36] N.E. Huang, Z. Shen, S.R. Long, M.L.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C. Tung, and H.H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," Proc. R. Soc. Lond. A, vol.454, no.1971, pp.903–995, March 1998
[37] M.A. Colominas, G. Schlotthauer, and M.E. Torres, "Improve complete ensemble EMD: A suitable tool for biomedical signal processing," Biomedical Signal Processing and Control, vol.14, pp.19–29, Nov. 2014.

Yaokun Hu received his B.E. degree in electrical engineering and automation from the College of Electric Information, Hunan Institute of Engineering, China, in 2017, and his M.E. degree in electrical engineering from the Graduate School of Science and Technology, Nihon University, Japan, in 2021, respectively. He is currently pursuing a Ph.D. degree in electrical engineering at the Graduate School of Science and Technology, Nihon University, Japan. He joined Fujitsu Ltd., Japan, in 2021, and is working in R&D. His current research interests include radar signal processing, biomedical signal processing, and machine learning.



Xuanyu Peng received her B.E. degree in electrical engineering and automation from the College of Electric Information, Hunan Institute of Engineering, China, in 2017, and her M.S. degree in electromechanical engineering from the Faculty of Science and Technology, University of Macau, Macau S.A.R, China, in 2020, respectively. From 2020 to 2023, she was an assistant teacher at the College of Intelligent Manufacturing, Hunan Vocational Institute of Technology, China. From 2023, She will be pursuing a Ph.D. degree in electrical engineering at the Graduate School of Science and Technology, Nihon University, Japan. Her current research interests include signal processing, control algorithms and intelligent robots.

Takeshi Toda received a B.E. degree in electrical engineering from Nihon University, Tokyo, Japan in 1992, an M.S. degree in electronic engineering from the University of Electro-Communications, Tokyo, Japan in 1994, and a D.E. degree from the Tokyo Institute of Technology, Tokyo, Japan in 2004. From 1994 to 2004, he worked at Fujitsu Laboratories Ltd., Kawasaki, Japan. From 2004 to 2005, he worked at eAccess Ltd., in Tokyo, Japan. From 2005 to 2008, he worked at the Kyocera Corp., R&D Center, Yokohama, Japan. He is currently a professor at the College of Science and Technology, Nihon University, Tokyo, Japan. His current research interests include radar signal processing, machine learning, and system-information engineering.