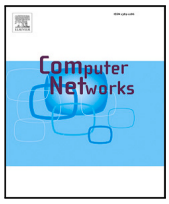
Computer Networks 257 (2025) 110989

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet



Coordinated multi-point by distributed hierarchical active inference with sensor feedback

Tatsuya Otoshi [a],*, Masayuki Murata [b]

[a] Graduate School of Economics, Osaka University, Osaka, Japan
[b] Graduate School of Information Science and Technology, Osaka University, Osaka, Japan

ARTICLE INFO

Keywords:
Active inference
Position estimation
Propagation channel estimation
Coordinated multi-point
Free energy principle

ABSTRACT

This study focuses on cooperative beamforming among base stations in wireless communication technology and proposes a new approach based on the Free Energy Principle (FEP). Traditionally, the trade-off between the accuracy of channel information acquisition and overhead has posed a challenge in beamforming. FEP addresses this trade-off by balancing the value from observation and the value from action to select the optimal behavior. This enables adaptive responses to dynamic environmental fluctuations through integrated search, reasoning, and learning. Additionally, by introducing a hierarchical coordination structure, information sharing among base stations and indirect state sharing among agents are achieved, enhancing the efficiency and stability of beamforming. Furthermore, this study utilizes Integrated Sensing and Communication (ISAC) to perform simultaneous communication and real-world sensing. By integrating feedback from terminals and terminal location information for channel state estimation, the overhead is reduced. Simulation results demonstrate that the proposed method effectively improves the Signal to Interference and Noise Ratio (SINR) and energy efficiency.

1. Introduction

In recent years, beamforming has been gaining importance in wireless communication technology. Beamforming is a technique to improve communication quality and capacity by controlling the directivity of transmitted radio waves using a transmitting antenna array. In particular, the combination of beamforming and Massive MIMO (Multiple Input Multiple Output) can significantly enhance directivity [1]. The use of a large number of transmit/receive antennas also enables simultaneous communication with multiple User Equipment (UE) devices.

However, in cellular communications consisting of multiple cells, coordination among base stations is required in the cell boundary region. In these regions, radio waves from multiple cells reach the terminal, resulting in significant interference effects. Therefore, there is a need for CoMP (Coordinated Multi-Point) operation, in which multiple base stations cooperate to perform beamforming and improve the communication quality of UE devices in the boundary region [2,3].

The effectiveness of the CoMP scheme depends on the accuracy of the channel state estimation. When accurate information is available, high throughput can be achieved by joint transmission (JT), where multiple base stations transmit the same signal [4]. Conversely, when accurate information is not available, cooperative scheduling/beamforming (CS/CB), in which multiple base stations transmit different signals while avoiding interference, is employed, albeit with reduced throughput compared to JT.

A critical challenge in cellular communications, especially in multi-cell environments, is the dynamic adaptation to the temporal variations of the Channel State Information (CSI) amidst environmental noise. The sources of noise include variations in channel state due to fading, feedback delays, quantization of feedback information, and hardware-induced noise. When beamforming is performed using erroneous feedback information, it may result in the selection of inappropriate beams or improper encoding and modulation [5]. This issue is exacerbated in the boundary regions of cells, where interference from multiple cells significantly affects signal quality. Hence, there is a pressing need for a robust solution that can adaptively respond to these environmental fluctuations and maintain high communication quality.

Simultaneously, the use of radio waves for both sensing and communication has been considered under the ISAC (Integrated Sensing and Communication) framework [6]. ISAC aims to utilize real-world information such as location data for channel state estimation, in addition to feedback from data communication. Although sensing and communication applications have been treated separately in previous ISAC studies and are often treated as simultaneous optimization of different

* Corresponding author.
E-mail address: t-otoshi@econ.osaka-u.ac.jp (T. Otoshi).

https://doi.org/10.1016/j.comnet.2024.110989
Received 1 September 2024; Received in revised form 3 December 2024; Accepted 9 December 2024
Available online 15 December 2024

objectives [7], this study aims to utilize sensing information to improve communication quality. However, sensed information also contains noise. For instance, an experiment with two 5G base stations estimated the position of a target using the Angle of Arrival (AoA) and evaluated its error [8]. In experiments targeting pedestrians, the average position estimation error was approximately 0.99 m, with temporary errors exceeding 3 meters during pedestrian tracking. Thus, integrating multimodal inputs for estimation and addressing the trade-offs in accuracy between modalities remain significant challenges.

To address these challenges, we draw inspiration from the human brain's functioning and employ the free energy principle (FEP) [9–11]. FEP aims to minimize free energy, a combination of search and use values. By minimizing free energy, FEP facilitates the selection of actions that balance exploration and exploitation. In [12], a method is proposed for beamforming using Deep Q-Network (DQN) across multiple base stations. However, the primary goal of reinforcement learning (RL) is utility maximization, and directly reducing the uncertainty of the system's perceived state is not its main objective. In contrast, the FEP considers both utility and uncertainty reduction, enabling the selection of optimal actions. FEP effectively integrates multiple information sources, such as location and feedback information, and adapts to environmental changes.

In this paper, we propose inter-cell cooperative beamforming using FEP. To achieve this, an FEP agent is placed in each base station, and a higher-level FEP agent coordinates their actions. The lower-level agents perform beamforming using information from the higher-level agents and feedback and sensing information from the UEs. The upper agents observe the states of the lower agents and provide feedback in the form of predicted states. This hierarchical structure enables cooperative beamforming and ensures consistency throughout the system. Additionally, we introduce a sensitivity adjustment mechanism within FEP, allowing the selection of appropriate actions even in environments with varying noise levels.

The novelty and contributions of this paper are as follows:

- We propose a method to improve communication quality by utilizing sensing information in pursuit of the true integration of communication and sensing information within ISAC. By employing the Free Energy Principle (FEP), this framework not only enhances communication quality but also incorporates the objective of reducing uncertainty, which is a fundamental goal of sensing.
- FEP enables the active selection of beams to acquire new information. In contrast, conventional methods such as DQN primarily focus on exploiting existing information, with action selection for acquiring new information relying on ad-hoc techniques like ϵ-greedy.
- We propose a hierarchical structure in which multiple FEP agents are deployed to coordinate beamforming across base stations. Lower-level agents perform beamforming using feedback and sensing information, while upper-level agents integrate and share information across the lower-level agents to ensure system-wide consistency.
- A sensitivity adjustment mechanism is introduced to handle fluctuations in environmental noise, allowing for the selection of appropriate actions even in environments where noise levels themselves vary.

The structure of this paper thereafter is as follows. In Section 2, we describe the system model that we treat in this paper. In Section 3, we give an overview of the active inference based on the free energy principle. Section 4 describes the beamforming method for multi-cell coordination using active inference. In Section 5, we describe the simulation results of the proposed method. In Section 6, we discuss the effectiveness and limitations of the proposed method. Section 7 presents a summary of this paper.

2. System model

2.1. Components

There are M macrocells and S small cells communicating with D UEs. In downlink communication, we assume a situation where radio waves from the base station by beamforming interfere with each UE.

The set of base stations $B^{(m)}$ of a macrocell, the set of small cells $B^{(s)}$, and the set of UEs U are defined as follows:

$$B^{(m)} = \{b_i^{(m)} | i = 1, \dots, M\} \tag{1}$$

$$B^{(s)} = \{b_i^{(s)} | i = 1, \dots, S\} \tag{2}$$

$$U = \{u_i | i = 1, \dots, D\} \tag{3}$$

Also, let $B = B^{(m)} \cup B^{(s)}$ be the set of all base stations.

2.2. Channel coefficients and beamforming

Each base station B transmits radio waves using M_b antennas, which are transformed by spatial characteristics to reach M_u antennas of the receiving UE u. The spatial characteristics are represented by the matrix $H^{u,b}(t)$, where the element $H_{ij}^{u,b}(t)$ represents the channel coefficients for transmission from antenna i to antenna j.

At each time, each UE u is connected to one of the base stations b, and the correspondence is denoted by $A(t)$. The element $a_{u,b}(t)$ of $A(t)$ is 1 if u and b are connected, and 0 otherwise.

Each base station performs beamforming by setting the phase and amplitude of the radio wave to be transmitted for each antenna. The phase is represented by the beam vector $\boldsymbol{w}_b(t)$ and the amplitude by $P_b(t)$. The following relationship is established between the transmitted signal $x_b(t)$ and the received signal $y_{u,b}(t)$:

$$\boldsymbol{y}_{u,b}(t) = \sqrt{P_b(t)} H^{u,b}(t) \boldsymbol{w}_b(t) \circ \boldsymbol{x}_b(t) \tag{4}$$

where $\circ$ denotes element-wise multiplication.

2.3. Radio interference

Since radio waves and noisy signals transmitted from other base stations reach UE u, the received signal at u is expressed as follows:

$$\boldsymbol{y}_u(t) = \boldsymbol{y}_{u,b}(t) + \sum_{b' \in B \setminus \{b\}} \boldsymbol{y}_{b'}(t) + \sigma_u(t). \tag{5}$$

At this time, SINR (Signal to Interference and Noise Ratio) is expressed as follows:

$$\gamma_u(t) = \frac{P_b(t) \| H^{u,b}(t) \boldsymbol{w}^b(t) \|^2}{I_u(t) + \sigma_u^2(t)} \tag{6}$$

$$I_u(t) = \left\| \sum\nolimits_{b' \in B \setminus \{b\}} \boldsymbol{y}_{b'}(t) \right\|^2 \tag{7}$$

where $I_u(t)$ is the strength of the interfering radio wave and $\sigma_u^2(t)$ is the strength of the noise.

2.4. Base station observation

The base station b is assumed to be able to observe the SINR $\boldsymbol{\gamma}_u(t)$ to each UE based on feedback from the terminal. However, this feedback may contain noise due to delays or incorrect feedback. The noisy observation $\tilde{\gamma}_u(t)$ is modeled as a Gaussian distribution with the true SINR $\gamma_u(t)$ as the mean:

$$\tilde{\gamma}_u(t) \sim \mathcal{N}(\gamma_u(t), \sigma^2) \tag{8}$$

where σ^2 represents the variance of the noise. The variance may change depending on the level of feedback error or environmental factors, influencing the observation accuracy.

2.5. Objective function

The transmission rate of data to UE u is estimated using SINR as follows [12,13]:

$$C_u(t) = \log(1 + \gamma_u(t)). \tag{9}$$

Therefore, the transmission rate can be maximized by maximizing SINR. Thus, maximizing SINR as one of the objective functions can improve beamforming performance. In FEP, however, the objective function is defined through a selection distribution. Details are given in Section 4.

Other objective functions may be used, such as power efficiency [12] or fairness-aware power efficiency [13]. Power efficiency is defined as the transmission rate per unit power consumption as:

$$EE = \frac{1}{T}\sum_{t=1}^{T}\frac{\sum_m \log(1+\gamma_m(t))}{P_{total}(t)} \tag{10}$$

where $\gamma_m(t)$ is the SINR of terminal m at time t and $P_{total}(t) = \sum_b (P_b(t) + P_{BS} + P_{UT})$ is the total power consumption at time t. $P_{BS} = 9$ dBW, $P_{UT} = 10$ dBm is the power dissipation at the base station and terminal.

In the method proposed in this paper, it is possible to use not only SINR but also power efficiency as an objective function. Therefore, in Section 5, we evaluate the case where power efficiency is also used as the objective function in addition to SINR.

2.6. Location sensing

In ISAC, the base station transmits signals for sensing and receives reflected waves from objects in the environment to estimate the location of the reflected objects [6,14,15]. In this case, the reflective object is not limited to a terminal, so it cannot be used alone to identify the terminal. Therefore, when multiple terminals are present, the location of all but the target terminal is also given without distinction.

Therefore, as location information, information on the density of reflective objects at each location is assumed to be obtained. Let $\rho(A)$ be the density of reflective objects in area A and ρ be the density of reflective objects in the whole area. Then ρ is defined as follows:

$$\boldsymbol{\rho} = \big(\rho(A_1), \ldots, \rho(A_K)\big) \tag{11}$$

where K is the number of areas and A_i represents the ith area.

With ISAC, it is possible to simultaneously acquire location information while maintaining the conventional communication rate. Although it is possible to acquire location information with higher accuracy by sacrificing part of the communication rate, this paper assumes that location information is acquired while maintaining the communication rate. In addition, it is assumed that the error in location information that occurs in this case is absorbed by the granularity of the area.

3. Active inference

In the context of active reasoning based on the free energy principle [11], the policy $\pi = (a_1, \ldots, a_T)$, representing the chosen sequence of actions over a period from time 1 to T, is determined to minimize the expected free energy over the entire time horizon. The expected free energy for the whole period is defined as:

$$\begin{aligned} G(o_{1:T}, s_{1:T}, \pi) \;=\; & E_Q[\log Q(s_{1:T}, \pi) \\ & - \log \tilde{P}(o_{1:T}, s_{1:T}, \pi)] \end{aligned} \tag{12}$$

where $o_{1:T}$ denotes observed values, $s_{1:T}$ denotes states, Q denotes the approximate posterior distribution, and $\tilde{P}$ denotes the target distribution. This formulation accounts for the expected free energy accumulated over the entire period, incorporating all observations and states from time 1 to T. The specific correspondence with beamforming, which is the subject of this paper, is discussed in Section 4.

At each specific time step τ, the expected free energy can also be calculated independently for that particular point in time. The expected free energy at a single time step τ is given by:

$$G_\tau(\pi) = E_{Q(o_\tau, s_\tau|\pi)}[\log Q(s_\tau|\pi) - \log \tilde{P}(o_\tau, s_\tau|\pi)] \tag{13}$$

$$\geq -E_{Q(o_\tau|\pi)}[D_{KL}(Q(s_\tau|o_\tau, \pi) \parallel Q(s_\tau|\pi))] - E_{Q(o_\tau|\pi)}[\log \tilde{P}(o_\tau)]. \tag{14}$$

This expression focuses on the expected free energy at a single time step τ, considering only the observations and states at that specific point in time. The first term in the above equation is the information gain of updating the posterior distribution by obtaining a new observed value o_τ. The second term represents the expected utility based on the observed value o_τ. This utility must be specified a priori by the prior distribution $P(o_\tau)$. Both the first and second terms are negative, and the expected free energy decreases as the information gain and utility increase.

In the case of beamforming with multiple base stations, the number of observations, states, and actions increases with the number of base stations. In particular, for states and actions, it is not realistic to minimize the expected free energy of the entire system since the distribution calculation must take into account their combinations. Therefore, it is desirable for the active inference agents placed at each base station to determine local actions based on local information. Each base station $b \in B$ determines a local action sequence $\pi^{(b)} = (a_1^{(b)}, \ldots, a_T^{(b)})$ that minimizes the following local free energy under the local observation information $o_\tau^{(b)}$ and local state $s_\tau^{(b)}$:

$$\begin{aligned} G_\tau(\pi)^{(b)} \;=\; & E_{Q(o_\tau^{(b)}, s_\tau^{(b)}|\pi^{(b)})}[\log Q(s_\tau^{(b)}|\pi^{(b)}) \\ & - \log \tilde{P}(o_\tau^{(b)}, s_\tau^{(b)}|\pi^{(b)})]. \end{aligned} \tag{15}$$

Coordination with other base stations is achieved by the exchange of local observation information $o_\tau^{(b)}$ and prediction/prediction errors. In FEP, coordination at the same hierarchical level is achieved by other agents observing the results of the agents' actions [16]. On the other hand, coordination between the upper and lower hierarchies is achieved through the exchange of predictions and prediction errors, where the upper hierarchy predicts the state of the lower hierarchy at a higher level of abstraction and the lower hierarchy communicates its prediction errors to the upper hierarchy [17,18].

4. Distributed beamforming with position and CSI feedback

In this section, we propose a method for distributed beamforming using active inference based on the Free Energy Principle (FEP). The overview of the proposed method is illustrated in Fig. 1. In the proposed method, lower-layer agents are placed at base stations that directly transmit beams to UEs, while upper-layer agents handle coordination among the base stations. The figure depicts a scenario where lower-layer agents are placed in small cells, and upper-layer agents are placed in macro cells. Lower-layer agents determine the beams using feedback from the targeted terminals and sensed density information of neighboring terminals as observed values. Upper-layer agents facilitate coordination by setting the expected states for each lower-layer agent based on the states estimated by the lower-layer agents.

The following subsections provide a detailed explanation of the lower-layer and upper-layer agents.

4.1. Lower layer agents

Lower layer agents are placed in each cell and determine the beam using feedback from terminals and sensed density information as observed values. Although the agents in the lower layer are independent of each other, they receive predictive distributions from the upper layer and cooperate with each other by realizing control according to their preference distributions. The detailed steps are outlined in Algorithm

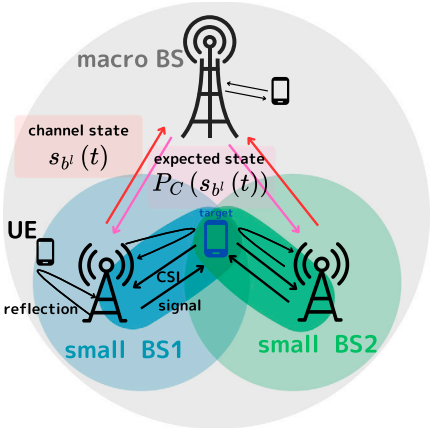


Fig. 1. Overview of the hierarchical coordination with active inference. Lower-layer agents are placed in small cells, and upper-layer agents are placed in macro cells. Lower-layer agents determine the beam using feedback from terminals and sensed density information as observed values. Upper-layer agents share the expected states for each lower-layer agent.

1, which shows how lower-layer agents operate in real-time, from observing SINR and estimating the channel coefficient, to determining the beam vector and transmit power.

Algorithm 1 Lower Layer Agent Operation at Each Time Step

Input: Feedback from UEs $\boldsymbol{\gamma}_b(t)$, Sensed density information $\boldsymbol{\rho}$, Predicted distribution from upper layer $\tilde{P}_C(s_b(t))$
Output: Beam vector $\boldsymbol{w}_b(t)$, Transmit power $P_b(t)$
1: Estimate system state $s_b(t)$ using Bayesian inference, where the prior distribution $\tilde{P}_C(s_b(t))$ is provided by the upper layer
2: Update internal model $P(s_{t+1}|s_t)$ and $P(o_t|s_t)$ through Bayesian inference
3: Infer the policy π that minimizes the expected free energy $G_t(\pi)$
4: Determine action $a_b(t) = (\boldsymbol{w}_b(t), P_b(t))$ with the policy π
5: Transmit beam with power $P_b(t)$ to UEs

4.1.1. Observation value

The base station b is assumed to be able to observe the SINR $\boldsymbol{\gamma}_b(t)$ to each UE based on feedback from the terminal.

At the same time as communicating data with the terminal, the base station also transmits a signal for sensing and receives its reflected wave to obtain the location information of a person in the environment. However, in consideration of the fact that it does not lead to the identification of a person, the acquired position of a person shall be observed as the density $\rho(A)$ for each area. Since the error in the estimated position is absorbed by dividing into areas, the uncertainty in the position is assumed to be mainly due to person-specific uncertainty.

4.1.2. State

Based on the observed SINR $\boldsymbol{\gamma}_b(t)$ and its own beamforming information, the channel coefficient $H^{u,b}(t)$ is estimated.

Since the area A where the targeted terminal is located is unknown, A is estimated based on density information and feedback from the terminal.

Thus, the state of the system corresponds to $s_b(t) = (H^{u,b}(t), \boldsymbol{\gamma}_b(t), A)$.

The prior distribution of states is given as the predicted distribution of the states of the lower layers by the upper layers, and the lower layers control themselves to conform to the assumptions of the upper layers, thereby realizing cooperation among the lower layers. Therefore, the prior distribution of the state in the lower layers is given by:

$$\tilde{P}_C(s_b(t)) = P(s_b(t)|o_{b^h}(t)) \tag{16}$$

where b^h denotes the macrocell base station.

4.1.3. Behavior

The base station specifies the beam shape by the beam vector $\boldsymbol{w}_b(t)$ and determines the transmit power $P_b(t)$. Thus, $a_b(t) = (\boldsymbol{w}_b(t), P_b(t))$ corresponds to the action.

4.1.4. Preference distribution

The preference distribution sets the desired state for the control as a prior distribution over the observed values.

In existing beamforming methods, the objective function is the power efficiency in Eq. (10) or transmission rate by maximizing SINR in Eq. (6). When setting such a specific objective function $R_b(o_\tau)$, since utility is measured by the logarithm of the preference distribution, the following Boltzmann distribution can be used to reflect $R(o_\tau)$ as the preference:

$$\tilde{P}_R(o_\tau) \propto \exp(\beta R(o_\tau)) \tag{17}$$

where β is a parameter that indicates how much the preference is biased by the size of the objective function. In this paper, simulation evaluation is performed using SINR or power efficiency as the objective function.

4.2. Upper layer agents

Upper layer agents are placed in macrocells, and the upper layer determines the CoMP setting (the group of lower layer agents that perform JT) using the states estimated by the lower layer agents as observed values. The predicted distribution of the upper layer states is used as a prior distribution for the lower layer states. Algorithm 2 illustrates the step-by-step process for upper-layer agents, from observing the states of lower-layer agents to determining the CoMP group and sending the predicted distributions back to lower-layer agents.

Algorithm 2 Upper Layer Agent Operation at Each Time Step

Input: States estimated by lower layer agents $P(s_{b^l}(t)|o_{b^l}(t))$
Output: CoMP group G, Predicted distribution for lower layer agents $\tilde{P}_C(s_b(t))$
1: Observe state $P(s_{b^l}(t)|o_{b^l}(t))$ from each lower layer agent b^l
2: Estimate system state $S_B(t)$ using Bayesian inference
3: Update internal model $P(S_{B(t+1)}|S_B(t))$ and $P(o_t|S_B(t))$ through Bayesian inference
4: Infer the policy π that minimizes the expected free energy $G_t(\pi)$
5: Select CoMP group G with the policy π
6: Send predicted distribution of states $\tilde{P}_C(s_b(t))$ to lower layer agents

4.2.1. Observation value

The estimated result $P(s_{b^l}(t)|o_{b^l}(t))$ of the state at the lower layer base station b^l is the observed value. Thus, it corresponds to $o_b(t) = P(s_{b^l}(t)|o_{b^l}(t))$.

4.2.2. State

The upper layer controls based on the channel coefficients $H^{u,b^l}(t)$ of each lower layer and the area A of the terminal being targeted. Thus, $S_B(t) = (A, H^{u,b^l_1}(t), H^{u,b^l_2}(t), \ldots)$.

4.2.3. Action

The upper layer determines the group of lower layer base stations that will perform JT. That is, we determine $a_b(t) = G$, where $G \subseteq B^{(s)}$ is a subset of the set of small cells.

4.2.4. Preference distribution

The preference distribution is assumed to be the Boltzmann distribution with power efficiency and transmission rate as in the lower layer. However, the upper layer differs from the lower layer in that it uses the aggregated power efficiency and transmission rate for the entire system.

4.3. Noise-resilient active inference

The Free Energy Principle estimates the magnitude of the error as an internal model. When the error is large, the influence of the error can be suppressed by reducing the weight of the observed values in the state update. However, for this to work appropriately, the internal model must be correctly learned. Nevertheless, it is considered difficult to estimate errors from the internal model, especially when there is a discrepancy between the observed values predicted by the internal model and the actual observed values. This is because it is challenging to distinguish whether the discrepancy is due to an error in the observed values or an error in the estimated state.

As a countermeasure, it is possible to distinguish between changes in the state of the external environment and a decrease in observation accuracy. In this case, since position information and radio wave strength feedback are used as observed values, it can be considered that if the prediction error occurs in conjunction with these, it indicates a change in the state of the external environment. Conversely, if the prediction error occurs independently, it indicates a decrease in observation accuracy. Then, if a decrease in observation accuracy occurs, the influence of the error can be suppressed by reducing the weight of the observed values.

Observation accuracy is represented by matrix A in the internal model. Matrix A represents the probability $A^{(m)}_{o,s}$ of obtaining the observed value o in state s for each modality m. A state with low observation accuracy is one in which $A^{(m)}_{o,s}$ varies for each observed value o when state s is fixed, and in the case of the lowest accuracy, it becomes a uniform distribution with respect to o.

Therefore, by approximating A to a uniform distribution as follows, it is possible to reflect the decrease in observation accuracy in the internal model:

$$d_o = A^{(m)}_{o,s} - \frac{1}{n} \tag{18}$$

$$a_o = A^{(m)}_{o,s} - \phi \cdot d_o \tag{19}$$

$$A'^{(m)}_{o,s} = \frac{a_o}{\sum_{j=1}^{n} a_o} \tag{20}$$

where, $\phi \in [0, 1]$ is a parameter representing the degree of decrease in observation accuracy, and the larger ϕ is, the greater the decrease in observation accuracy.

5. Evaluation

Simulations are performed to confirm the effectiveness of using ISAC position estimation in beamforming with longitudinal coordination by FEP.

5.1. Simulation setting

The simulation environment is designed to match the setting in [12], allowing for a direct comparison of energy efficiency between the proposed method and the baseline method. For one UE with one antenna, two base stations with $N = 4$ antennas transmit radio waves. To check the effectiveness of the location estimation, we evaluate it in a situation where the terminal is moving. The terminal moves in a random walk. The random walk is assumed to follow a Gaussian distribution with a standard deviation of 1 m per second. The speed converted to mph is 3.6 km/h, which is almost the same as the walking speed of a person. Fig. 2 shows the trajectory of a terminal on a random walk. The triangles in the diagram indicate the location of the base station. Other parameters are shown in Table 1. Transmit power $P_b(t)$ is chosen from 5 levels from 0 to $P_{\max}$ where $P_{\max} = 38$ dBm.

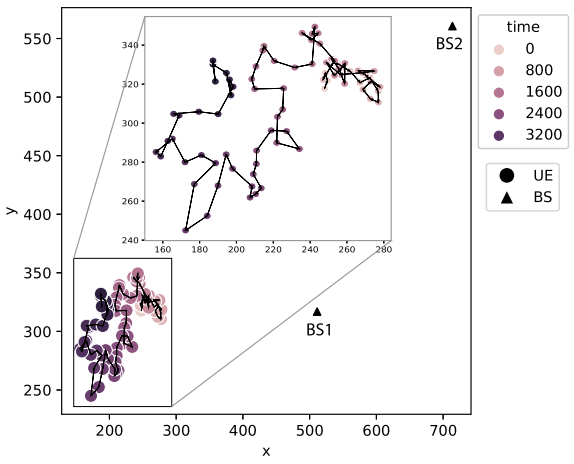


Fig. 2. Trajectory of terminal movement and base station location.

Table 1
Evaluation environment.

Parameters	Simulation values
Cell radius	200 m
Noise intensity	−114 dBm
Maximum transmit power	38 dBm
Power levels	5 levels
Multipath spread of the transmission angle	3 degrees
Attenuation over distance	$120.9 + 37.6 \lg(d)$ dB

Depending on the location of the terminal, the channel coefficients are set as follows:

$$h_{ij} = \sqrt{\frac{\beta_{i,j}}{L}} \sum_{l=1}^{L} a^{\dagger}_{i,j}(\theta_l) \tag{21}$$

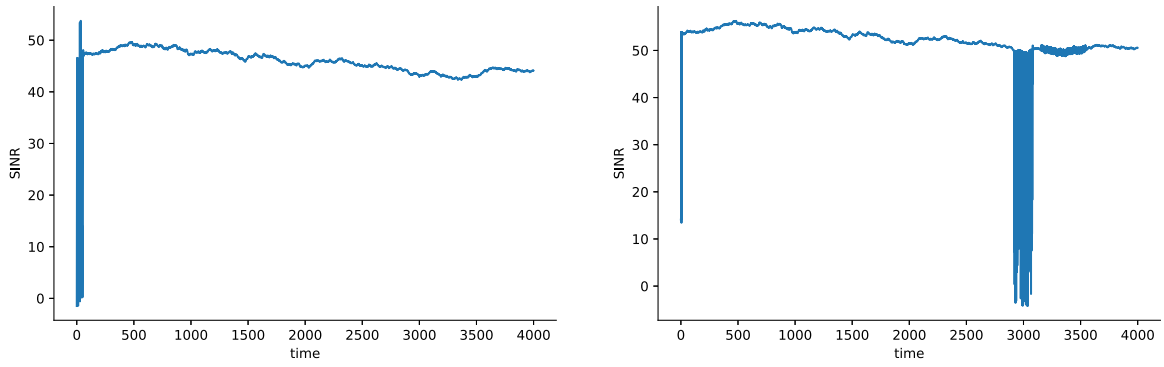
$$a_{i,j}(\theta_l) = \frac{1}{\sqrt{N}} \left(1, \exp(l\pi i \cdot 1 \cos\theta_l), \ldots, \exp(l\pi i \cdot (N-1)\cos\theta_l)\right) \tag{22}$$

where L is the number of paths from the base station to the terminal and we set $L = 4$. θ_l denotes the transmission angle of path L and follows a uniform distribution of $(\bar{\theta}_l - \vartheta/2, \bar{\theta}_l + \vartheta/2)$. The mean value $\bar{\theta}_l$ of the sending angle changes as the terminal moves. The $\beta_{i,j}$ represents the attenuation due to distance, which also changes according to the movement of the terminal. The noise intensity is assumed to be −114 dBm.

The beamforming vectors shall be selected from N types of pre-prepared codebooks. The nth beamforming vector is as follows:

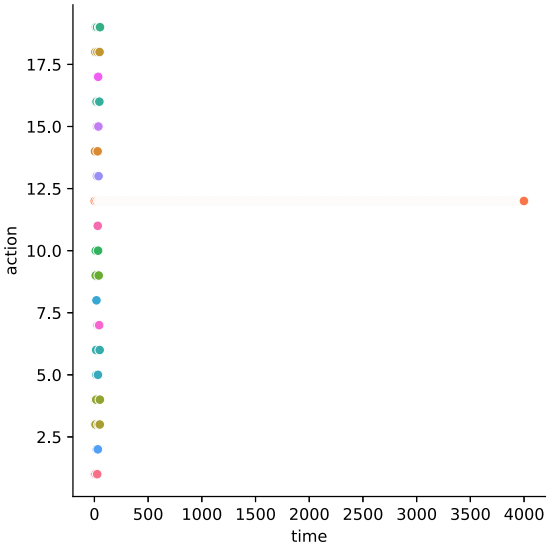
$$\boldsymbol{w}_n = \left(\frac{1}{\sqrt{N}} \exp\left(\frac{2\pi i(n-1)0}{N}\right), \ldots, \frac{1}{\sqrt{N}} \exp\left(\frac{2\pi i(n-1)(N-1)}{N}\right) \right). \tag{23}$$

The base station observes the SINR from the UEs as feedback and also observes the position of each UE as density information in space. When the observation results are obtained, the internal model and state are updated, and the power and beamforming vector of the next beam to be transmitted are determined. The initial values of the internal model $P(o_\tau|s_\tau, \pi)$ and $P(s_\tau|s_{\tau-1}, \pi)$ are set to random conditional probabilities as initial models.

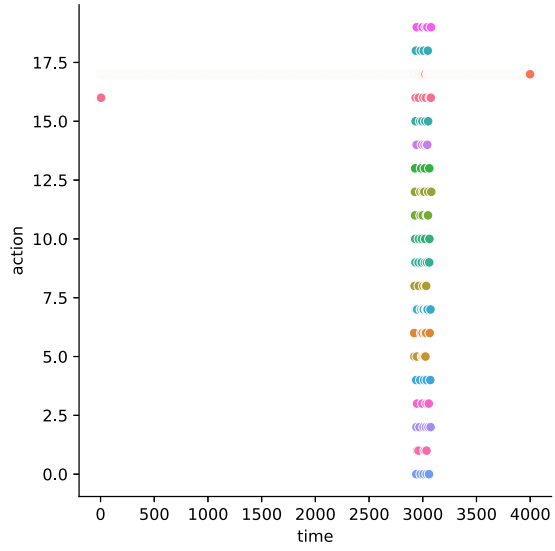


(a) SINR (feedback + location)

(b) SINR (feedback only)



(c) Time series of beams selected in cell 1 (feedback + location)



(d) Time series of beams selected in cell 1 (feedback only)

Fig. 3. Time series of SINR and selected beams with and without location information.

5.2. Results with maximizing SINR

To confirm the effect of using sensing by ISAC, we compare the behavior with and without location information (density information) as input to the FEP. First, we use SINR as the objective function.

Fig. 3 shows the simulation results with and without location information. (a) and (b) are the SINR time series for each case, and (c) and (d) represent the beams selected at each time in each case. The unit time on the x-axis corresponds to the time step, $\tau = 1$ s.

From the figure, it can be seen that the period of stable and high SINR is longer and the beam changes less when location information is used.

Around time 3000, the UE has moved away from its initial position, and the SINR is also moving towards a slight decrease. If only feedback information is used, the UE reacts to this decrease and starts searching for the beam. On the other hand, when position information is used, the change of beam is suppressed because it is known that the decrease in SINR is due to a change in position and should not be responded to by changing the direction of the beam. Thus, it was observed that the use of position information stabilizes the beam selection and keeps the SINR high.

In both cases, a temporarily low SINR is recorded immediately after the simulation starts, but this is due to the random initialization of the model. In practice, the model is iterated after training, so the initial drop is not an issue.

Fig. 4 shows the average SINR for different input information. The average SINR was calculated for 100 different movement patterns. In the figure, DM is the result when the position information is input as a density map, SINR is the result when SINR is input as feedback, and SINR+DM is the result when both are input.

The figure shows that by using feedback and location information together, a higher SINR is achieved than when each is used separately, indicating that the information is properly integrated.

5.3. Results with maximizing power efficiency

Fig. 5 shows the time variation of power efficiency with and without location information. From the figure, when location information is used, the power efficiency reaches 350, which is almost the same as the highest value of the power efficiency of the method proposed in Ref. [12]. It takes about 25,000 steps to reach the highest power efficiency in the [12] method, while the proposed method reaches the highest value in a short period of time. On the other hand, when no location information is used, the power efficiency reaches 350 temporarily but does not stabilize and settles near 250. The power efficiency of 250 is equivalent to the power efficiency of the case where each base station uses only its own power efficiency as the objective function in the literature [12], and is not the overall optimum, but the power efficiency when each cell maximizes its own power efficiency. When location information is not used, it is considered that due to the

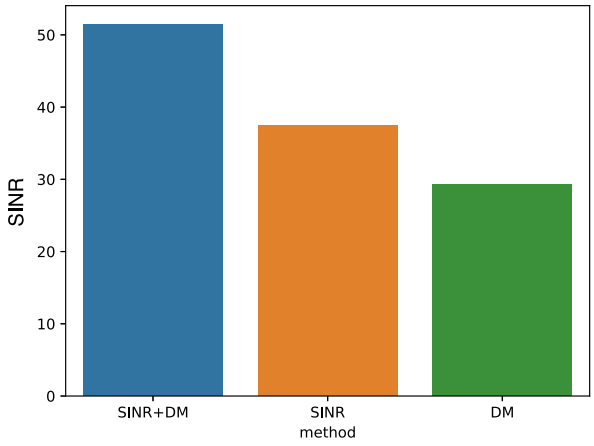


Fig. 4. Average SINR with different input information.

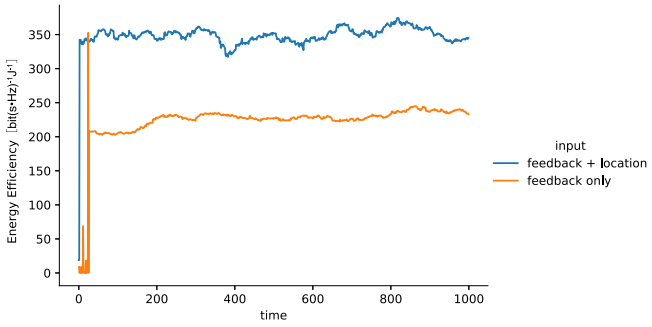


Fig. 5. Timeseries of the power efficiency.

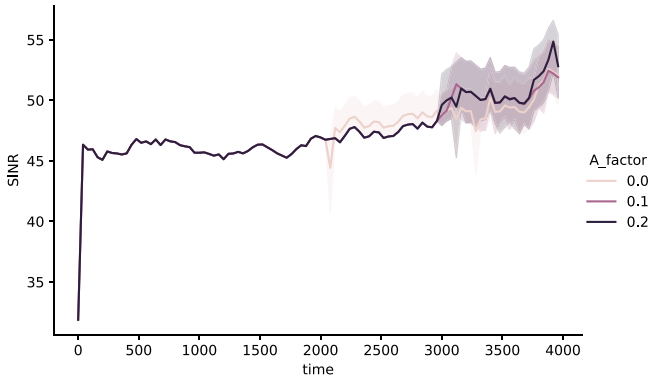


Fig. 6. Time series of SINR with different ϕ.

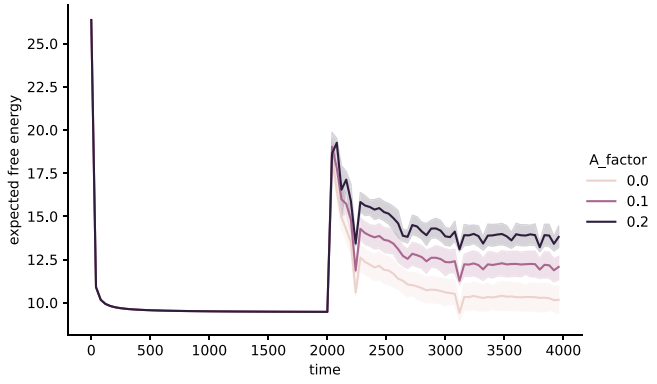


Fig. 7. Time series of expected free energy with different ϕ.

uncertainty of the channel state, coordination among cells cannot be achieved and we are stuck with a local solution. As shown above, the effectiveness of the method is demonstrated by the fact that the overall optimal beamforming can be achieved in a short period of time by using location information.

5.4. Results with noise-resilient active inference

To investigate the impact of observation errors on control and the effectiveness of Noise-Resilient Active Inference, simulations were conducted. The agent's observation information consists of two types: feedback information of channel state and position information, each considering its respective errors.

For the feedback of the channel state, observation information with errors is obtained by adding Gaussian noise to the SINR in Eq. (8). By changing the variance of the Gaussian noise, the degree of error is varied. For position information, observation values with errors are obtained by adding errors to the actual positions. Since position information pertains to the number of people per area, when errors cause misclassification across area boundaries, the terminals are considered to exist in the neighboring area a' instead of the actual area a. By changing the frequency of misclassifications to neighboring areas, the degree of error is varied.

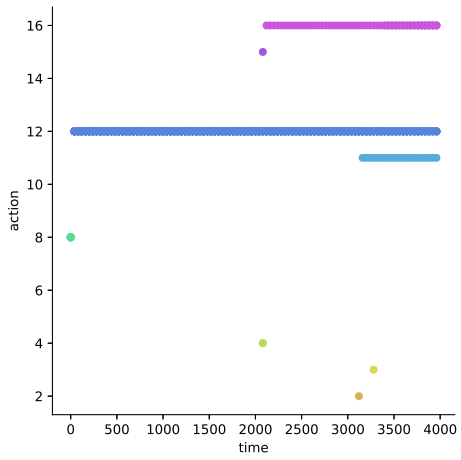
Fig. 6 shows the time series of SINR when the standard deviation of noise is changed during the simulation. At time 2000 in the simulation, the standard deviation of the noise is changed from 0 to 5, and the simulation continues with the constant standard deviation thereafter. The figure shows the time series of SINR when the parameter ϕ, which represents the degree of decline in observation accuracy, is varied. Additionally, Fig. 7 shows the time series of expected free energy, which represents the prediction error with inner model, and Fig. 8 shows the time series of the selected beams.

From the figures, it can be observed that when $\phi = 0$, i.e., no correction for observation accuracy in the internal model, beam switching starts at time 2000 when noise begins to be added. Initially, beams that temporarily lower the SINR are selected, but subsequently, slightly higher SINR beams are chosen. However, after time 3000, different beams are also selected, resulting in a generally lower average SINR compared to when corrections are made. The expected free energy change is also the fastest for $\phi = 0$, indicating that it is sensitive to noise.

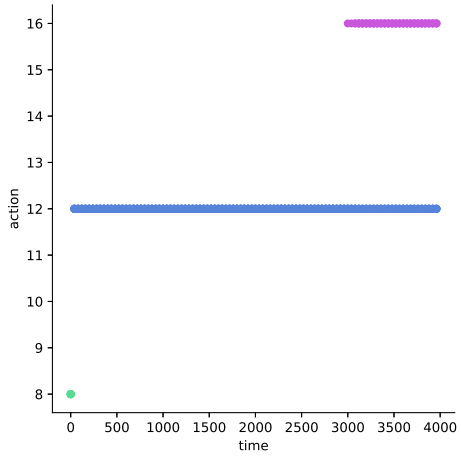
On the other hand, when observation accuracy corrections are made for $\phi = 0.1$ and 0.2, it is observed that the same beams continue to be selected for a while even when noise is added. This is because the corrections for observation accuracy suppress the impact of noise. Subsequently, beam switching occurs around time 3000, which is due to the user's movement causing a change in the appropriate beam. The beams selected are the same as those selected earlier in the case of $\phi = 0$, but in the case of $\phi = 0$, additional different beams are selected after time 3000 due to the influence of noise, resulting in a generally higher average SINR for $\phi = 0.1$ and 0.2 compared to $\phi = 0$. Increasing ϕ also slows down the change in expected free energy, indicating that changes in the internal model due to overreaction to noise are suppressed.

Thus, it was confirmed that improving the SINR is possible by suppressing hypersensitive reactions due to noise through corrections in observation accuracy.

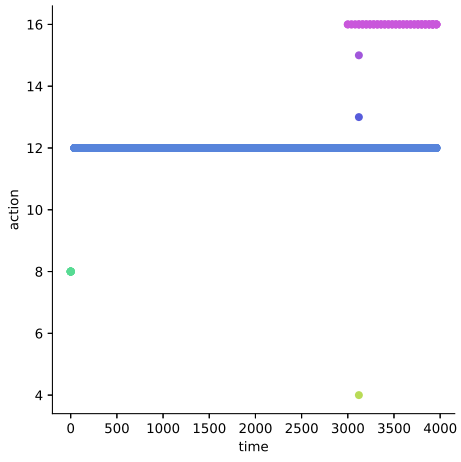
However, in all cases, SINR fluctuations due to continuous switching of multiple beams are observed. This occurs because the presence of



(a) ($\phi = 0$)



(b) ($\phi = 0.1$)



(c) ($\phi = 0.2$)

Fig. 8. Time series of selected beams with different ϕ.

noise in the observation values leads to exploration to reduce uncertainty. Since the gain from exploration is included in the objective function under the free energy principle, in the presence of noise, the relative gain from exploration increases, leading to frequent beam switching. This intuitively reasonable behavior, where exploration occurs in the presence of noise and a single beam is continuously selected in its absence, emerges as a result of minimizing free energy.

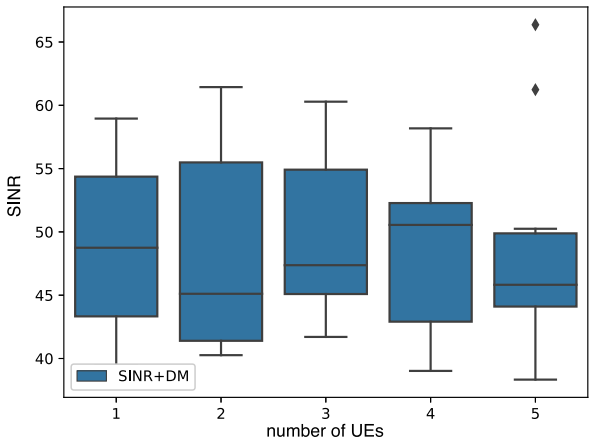


Fig. 9. SINR with different numbers of UEs. The box plot displays the distribution of SINR values for each group, where the central line indicates the median, the box represents the interquartile range (IQR), and the whiskers extend to the smallest and largest data points within 1.5 times the IQR from the quartiles. Outliers, shown as individual points, are defined as values beyond this range.

5.5. Scalability

To evaluate the scalability of the proposed method, we present simulation results with varying numbers of UEs. Since the positioning sensing treats the presence of UEs as a density distribution without distinguishing between individual UEs, increasing the number of UEs tends to degrade the accuracy of the target UE's position estimation. However, the proposed method also utilizes feedback from the target UE's received signal strength, allowing it to integrate this feedback with the density information. As a result, an increase in the number of UEs does not necessarily lead to a significant degradation in position estimation accuracy or communication performance.

Fig. 9 shows the box plot of SINR values as the number of UEs varies. It can be observed that the variation in SINR values across different numbers of UEs is comparable to the variation within each group of UEs. This suggests that the proposed method effectively integrates CSI feedback with position information, maintaining reliable position and channel estimation even as the number of UEs increases.

6. Discussion

The effectiveness of utilizing location information in beamforming has been demonstrated through the proposed approach. In contrast to conventional position-aided beamforming approaches [19] that rely on precise position information of a certain target UE, the proposed method leverages uncertain sensing data combined with CSI feedback. This innovative integration enables effective beamforming even in scenarios where the target UE cannot be distinctly identified, such as ISAC sensing. By integrating location information with CSI feedback, it becomes possible to achieve stable beamforming. This is particularly evident in the fact that the proposed method does not rely on directly measuring the position of the target UE, but instead utilizes the density distribution of multiple UEs and the integrated CSI feedback to derive information that can be used for beamforming. This dual approach enhances the system's adaptability and performance.

The key distinction between active inference and reinforcement learning is also notable. In traditional reinforcement learning [12],

the agent learns behavior through interaction with the environment, typically using methods like the ϵ-greedy algorithm for exploratory actions. However, in active inference, the value of exploration itself is embedded within the calculation of free energy, allowing the agent to actively explore based on its internal model when necessary. This method is inherently more suited to Integrated Sensing and Communication (ISAC) systems, as it not only aims to improve communication performance but also optimizes the sensing process. This dual optimization aligns well with the goals of ISAC, ensuring efficient resource usage and performance stability.

Scalability is another important aspect of the proposed method. When utilizing reflections of radio waves for position estimation, the position of a specific UE is not directly measured, and as the number of UEs increases, the accuracy of target UE position estimation can degrade. Nevertheless, the proposed approach has been shown to maintain communication performance even as the number of UEs increases. This is due to the ability to effectively integrate CSI feedback with location information, allowing for accurate position and channel estimation despite the increased number of UEs. In the present study, simulations were limited to two base stations, but the scalability of this approach to larger systems should be explored further. As the number of base stations increases, the amount of information aggregated by the upper-level base stations will also increase, potentially requiring a multi-tier hierarchical structure to handle the additional complexity efficiently.

In terms of dynamics, this study assumed a relatively slow-moving environment where UEs move at walking speed. The results show that the proposed method can appropriately adapt to changes in channel state under these conditions. However, for faster-moving UEs, where channel state changes more rapidly, the method may need to incorporate faster information sharing among lower-level base stations without relying solely on upper-level coordination. This could involve techniques such as low-latency communication between lower-level agents or decentralized decision-making processes to maintain optimal beamforming under more dynamic conditions.

The proposed method presents a novel approach to cooperative beamforming, integrating location sensing and communication in a way that balances information acquisition and control. While the results are promising, future work should focus on real-world experiments to validate the method in practical environments. Furthermore, expanding the scope of ISAC beyond location information to incorporate additional sensing modalities could further enhance system performance and broaden the applicability of the approach.

Additionally, this study assumes fixed base stations, without considering environments where both the base stations and the UEs are mobile, such as in mobile base stations or sidelink communication scenarios. In cases where both sides are moving, even if their relative positions remain constant, the radio environment may vary due to dynamic conditions. In such scenarios, it may be necessary to manage spatial information more dynamically, for instance, through geographic-based spatial information management to account for the changing radio environment.

7. Conclusion

This study proposes a method for balancing information acquisition and control in cooperative beamforming among multiple base stations by considering the trade-off between information acquisition and control performance. The method is based on the Free Energy Principle (FEP), where FEP agents installed in base stations perform beamforming using feedback from user terminals and sensing information. The coordination among agents improves the efficiency and stability of beamforming as a whole by indirectly sharing the agents' states through the introduction of a hierarchical structure. Simulation results demonstrate that by combining channel information from terminal feedback with location information sensed by ISAC, more stable beamforming can be achieved, leading to improvements in average SINR and energy efficiency. Furthermore, by introducing sensitivity adjustments within FEP, noise-resilient beamforming can be achieved even when the level of noise varies due to environmental changes.

Future work includes verifying the effectiveness of this method through experiments in real-world environments. Additionally, expanding the scope of ISAC sensing beyond human movement and location information to integrate a broader range of data could further enhance beamforming performance.

CRediT authorship contribution statement

Tatsuya Otoshi: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Masayuki Murata:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Tatsuya Otoshi reports financial support was provided by Japan Society for the Promotion of Science (JSPS).

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number JP24K20761.

Data availability

No data was used for the research described in the article.

References

[1] F.A.P. de Figueiredo, An overview of massive MIMO for 5G and 6G, IEEE Lat. Am. Trans. 20 (6) (2022) 931–940.
[2] M.S.J. Solaija, H. Salman, A.B. Kihero, M.İ. Sağlam, H. Arslan, Generalized coordinated multipoint framework for 5G and beyond, IEEE Access 9 (2021) 72499–72515.
[3] T.K. Tandra, F. Tajrian, A. Hossain, M.T. Kawser, M.R. Akram, A.B. Shams, Joint transmission coordinated multipoint on mobile users in 5G heterogeneous network, in: 2022 IEEE 2nd Conference on Information Technology and Data Science, CITDS, IEEE, Debrecen, Hungary, 2022, pp. 273–278.
[4] S. Kumagai, T. Kobayashi, T. Oyama, C. Akiyama, M. Tsutsui, D. Jitsukawa, T. Seyama, T. Dateki, H. Seki, M. Minowa, et al., Experimental trials of 5G ultra high-density distributed antenna systems, in: 2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall), IEEE, Honolulu, HI, USA, 2019, pp. 1–5.
[5] A.İ. Tunali, H.A. Çirpan, Impact of imperfect channel estimation on 5G-NR, in: 2021 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom), IEEE, virtual conference, 2021, pp. 1–6.
[6] A.M. Elbir, K.V. Mishra, S. Chatzinotas, M. Bennis, Terahertz-band integrated sensing and communications: Challenges and opportunities, 2022, arXiv preprint arXiv:2208.01235.
[7] N. Huang, H. Dong, C. Dou, Y. Wu, L. Qian, S. Ma, R. Lu, Edge intelligence oriented integrated sensing and communication: A multi-cell cooperative approach, IEEE Trans. Veh. Technol. (2024).
[8] R. Liu, M. Jian, D. Chen, X. Lin, Y. Cheng, W. Cheng, S. Chen, Integrated sensing and communication based outdoor multi-target detection, tracking, and localization in practical 5G networks, Intell. Converged Netw. 4 (3) (2023) 261–272.
[9] K. Friston, The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. 11 (2) (2010) 127–138.
[10] C. Heins, B. Millidge, D. Demekas, B. Klein, K. Friston, I. Couzin, A. Tschantz, Pymdp: A python library for active inference in discrete state spaces, 2022, arXiv preprint arXiv:2201.03904.
[11] L. Da Costa, T. Parr, N. Sajid, S. Veselic, V. Neacsu, K. Friston, Active inference on discrete state-spaces: A synthesis, J. Math. Psych. 99 (2020) 102447.
[12] K. Yu, G. Wu, S. Li, G.Y. Li, Local observations-based energy-efficient multi-cell beamforming via multi-agent reinforcement learning, J. Commun. Inf. Netw. 7 (2) (2022) 170–180.

[13] H. Al-Obiedollah, K. Cumanan, J. Thiyagalingam, A.G. Burr, Z. Ding, O.A. Dobre, Energy efficiency fairness beamforming designs for MISO NOMA systems, in: 2019 IEEE Wireless Communications and Networking Conference, WCNC, IEEE, Marrakech, Morocco, 2019, pp. 1–6.
[14] F. Liu, Y. Cui, C. Masouros, J. Xu, T.X. Han, Y.C. Eldar, S. Buzzi, Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond, IEEE J. Sel. Areas Commun. 40 (6) (2022) 1728–1767.
[15] C. Ouyang, Y. Liu, H. Yang, N. Al-Dhahir, Integrated sensing and communications: A mutual information-based framework, IEEE Commun. Mag. 61 (5) (2023) 26–32.
[16] J. Pöppel, S. Kahl, S. Kopp, Resonating minds—Emergent collaboration through hierarchical active inference, Cogn. Comput. 14 (2) (2022) 581–601.
[17] O. Çatal, T. Verbelen, T. Van de Maele, B. Dhoedt, A. Safron, Robot navigation as hierarchical active inference, Neural Netw. 142 (2021) 192–204.
[18] G. Pezzulo, F. Rigoli, K.J. Friston, Hierarchical active inference: a theory of motivated control, Trends Cogn. Sci. 22 (4) (2018) 294–306.
[19] T.-H. Chou, N. Michelusi, D.J. Love, J.V. Krogmeier, Fast position-aided MIMO beam training via noisy tensor completion, IEEE J. Sel. Top. Sign. Proces. 15 (3) (2021) 774–788.

Tatsuya Otoshi received an M.E. and D.E. degree in information science and technology in 2014 and 2017 from Osaka University, where he is an Assistant Professor with the Graduate School of Economics. His research interests include traffic engineering and traffic prediction. He is a student member of the IEEE.

Masayuki Murata received the M.E. and D.E. degrees in information and computer science from Osaka University, Japan, in 1984 and 1988, respectively. Since April 2004, he has been a Professor with the Graduate School of Information Science and Technology, Osaka University. His research interests include information network architecture, performance modeling, and evaluation. Prof. Murata is a member of ACM and IEICE.