Cell Association for mmWave/Thz Ultra-dense Network using Computer Vision

Khoa Anh Ngo, Hyunsoo Kim, and Byonghyo Shim

Seoul National University.

{ngoak, hskim}@islab.snu.ac.kr, bshim@snu.ac.kr

Abstract

Ultra-dense networks (UDN) are anticipated to fulfill the high-performance demands for future applications of 5G and beyond communication networks. In this paper, we propose a novel cell association method that exploits the recent advances in computer vision techniques. The computer vision technique provides 3D position and the line-of-sight condition of mobile users within the UDN using RGB images. From the simulations, we demonstrate that the proposed outperforms 5G-NR in terms of total data rate.

Ⅰ. Introduction

Millimeter-wave (mmWave) and terahertz (THz) ultra-dense networks (UDNs) have been proposed as pivotal technology to meet the demands of immersive and interactive applications in 5G and beyond networks [1]. An ultra-dense network (UDN) deploys a large number of small base stations (SBSs) to circumvent the severe attenuation of mmWave/THz band signals [2]. It is in the interest of exploiting the benefits of UDN that we should associate mobile users with SBSs that facilitate line-of-sight (LoS) communication.

An aim of this paper is to propose a novel cell association technique based on the sensing and computer vision (CV). That is, we exploit 3D object detection, a CV technique that provides the 3D position of objects, to determine the beamforming angles of mobile users. In addition, since the sensing devices (e.g., RGB cameras, and infrared cameras) capture the visual information of visible objects, the appearance of mobile users can indicate the LoS conditions. Using the spatial information and LoS conditions, the proposed technique can associated mobile users with SBSs that maximize the data rate.

Ⅱ. Method

The main goal of the proposed technique is to exploit the spatial information and LoS conditions of UEs extracted from visual sensing data (e.g., RGB images) to generate the cell association vector $\hat{\boldsymbol{c}}_k \in \{0,1\}^M$, where M is the number of SBSs. To reliably detect and extract spatial information of all UEs, we perform the voxelization process that divides the 3D space of the UDN into fixed-size boxes called voxels. By allocating visual features corresponding to each voxel, we can exploit the detailed texture and 3D structure for reliable UE detection.

Ⅱ.1. Multi-view-based 3D object detection

In this subsection, we describe the operations of the multi- view-based 3D object detection that consists of three main components: 1) feature extractor, 2) voxelizer, and 3) head.

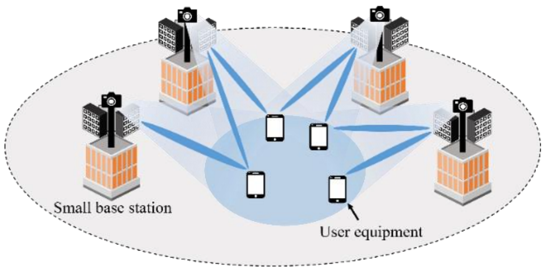

Fig. 1: Computer vision-aided ultra-dense network.

1. **Feature extractor:** A feature extractor combines the local features (e.g., shape and texture) and global features (e.g., face and wall) from the input RGB images into multi-scale features F_m.
2. **Voxelizer:** A voxelizer allocates the averaged bilinearly sampled visual features to the voxel at position $[x\ y\ z]$, $\tilde{\boldsymbol{q}}_{[x\,y\,z]} = 1/M \sum_{m=1}^{M} f_{bilinear}(F_m,\ [w\ h]_m^T)$, where $[w\ h]_m^T$ is the projected pixel of the voxel on the m-th camera coordinate system. While different objects have contradicting visual features, e.g., different colors, or shapes, aggregating visual features of the same object would result in detailed texture and 3D structure, effectively indicating the LoS conditions as blocking objects would have different visual information than UEs.
3. **Head:** Finally, a head generate a class score $\hat{\boldsymbol{s}}_{[x\,y\,z]} \in [0,1]^C$, where C is the number of classes that are designated as a mobile user, e.g., phone, tablet, laptop. We select $\hat{K}$ voxels, whose class score $\hat{\boldsymbol{s}}_{[x\,y\,z]_k}$ is higher than a pre-defined threshold $\delta \in [0,1]^C$ as UEs. For brevity, we denote the features of k-th UE voxel at position $[x\ y\ z]_k^T$ as $\boldsymbol{q}_k \coloneqq \tilde{\boldsymbol{q}}_{[x\,y\,z]_k}$. Then, the head obtains the precise position of the k-th UE $\hat{\boldsymbol{p}}_k = [x\ y\ z]_k^T + \boldsymbol{\Delta}_k$, where Δ_k is a generated offset based on the k-th voxel features.

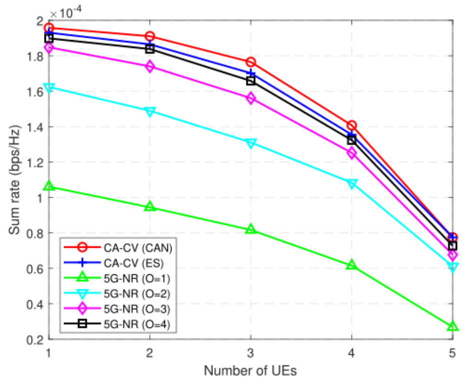

Fig 3: Sum rate as a function of the number of UEs.

II.2. Cell association network

We associate mobile users with SBSs using a cell association network (CAN) based on the constructed voxels. Basically, we view the cell association problem as a binary classification problem generating a cell association vector $\hat{\boldsymbol{c}}_k \in \{0,1\}^M$. To satisfy the rate requirement of the k-th UE R_k^{min}, we convert the concatenation of k-th voxel features and k-th rate requirement $[\boldsymbol{q}_k\ R_k^{min}]$ to $\boldsymbol{g}_k$ using a learnable linear embedding $W_{in} \in R^{(D+1)\times D}$:

$$\boldsymbol{g}_k = W^{in}[\boldsymbol{q}_k\ R_k^{min}].$$

After input embedding, the inputs are passed through CAN that consists of two fully connected layers with a non-linear activation, i.e., ReLU, in between:

$$\hat{\boldsymbol{c}}_k = f_{sigmoid}(f_{ReLU}(W_1^{FFN}\boldsymbol{g}_k + \boldsymbol{b}_1^{FFN})\boldsymbol{W}_2^{FFN} + \boldsymbol{b}_2^{FFN}),$$

where $f_{sigmoid}(x) = 1/(1+e^{-x})$, $f_{ReLU}(x) = \max(0,x)$, and $\boldsymbol{W}_1^{FFN}, \boldsymbol{W}_2^{FFN}, \boldsymbol{b}_1^{FFN}, \boldsymbol{b}_2^{FFN}$ are learnable weights and biases. Finally, the MBS sends the positions of associated UEs to the corresponding SBSs. Each SBS transmits a minimum number of pilot signals to establish communication.

III. Simulations

In this section, we examine the performance of the proposed cell association technique in terms of positioning error and data rate.

III.1. Simulations setup

We investigate the performance of the proposed cell association technique in terms of positioning error and data rate. To evaluate the 3D object detector and CAN, we generate datasets utilizing a 3D rendering engine and report the throughput of the proposed and benchmark scheme. We investigate the performance of CV-CA. We consider the UDN scenario where $M = 4$ SBSs cooperatively serve K users uniformly distributed within a square of $17 \times 22\,m^2$. The rate requirement of each user is uniformly chosen from $[0, R_{req}^{max}]$. The small-scale fading coefficients are generated according to the complex Gaussian distribution $\alpha_{m,k} \sim CN(0,1)$ and the large-scale fading coefficient is given by $\beta_{m,k} = \left(\frac{c}{4\pi f_c d_{m,k}}\right)^2$ where c is the speed of light, $f_c = 100\,Ghz$ is the carrier frequency, and $d_{m,k}$ is the distance between m-th SBS and k-th UE. We

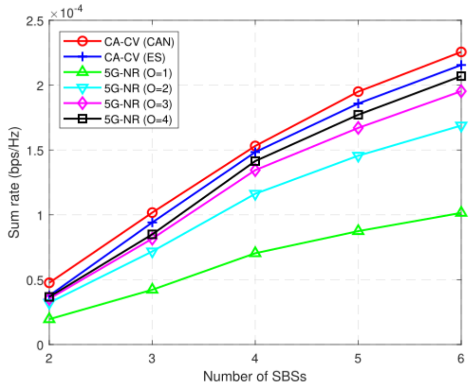

Fig 2: Sum rate as a function of the number of SBSs.

adjust the transmission power to maintain the SNR of 20dBm. We generate a synthetic dataset consisting of multi-view RGB images. For each sample, we obtain labels including the position of UEs and SBSs, and LoS conditions. We synthesized 500 and 250 samples for training and testing. We compare the proposed technique with 5G NR and an exhaustive search.

III.1. Simulations

In Table. 1, we summarize the positioning error and angle error. We observe that the proposed technique estimated the position of UEs with high accuracy, achieving the average error of 7cm. Whereas, in the codebook-based beam sweeping, the mismatch between the pre-defined beam direction and the real direction is unavoidable, resulting in an average angle error of $5.6°$.

In Fig. 2, we plot the average throughput as a function of the number of users K. We observe that the proposed CV-CA outperforms the conventional strategies in all scenarios. Specifically, the proposed CV-CA obtains 90% gain in throughput and 80% improvement in latency reduction over the conventional RSRP-based cell association technique. Additionally, in Fig. 3, we plot the average throughput as a function of the number of SBSs. We observe that the proposed CV-CA achieves the best average sum rate at any number of SBSs. We contribute the throughput gain of the proposed CV-CA to the sharp beamforming angles.

ACKNOWLEDGMENT

This work was supported by the Institute of Information and Communications Technology Planning and Evaluation (IITP) Grant funded by the Korean Government (MSIT, Development of Intelligent Wireless Access Technologies) under Grant 2021-0-00972.

REFERENCES

[1] H. Lee, B. Lee, H. Yang, J. Kim, S. Kim, W. Shim, B. Shim, and H. V. Poor, "Towards 6G hyper-connectivity: Vision, challenges, and key enabling technologies," J. Commun. Net., 2023.

[2] A.-A. A. Boulogeorgos, S. K. Goudos, and A. Alexiou, "User association in ultra dense Thz networks," in 2018 IEEE 19th Int. Work. Sig. Proc. Adv. in Wireless Commun. (SPAWC), 2018, pp. 1-5.