Pilot Beamforming for Massive Machine Type Communication Systems

Hyoungju Ji*, Jun Won Choi† and Byonghyo Shim*

* INMC, Dept. of Electrical and Computer Engineering, Seoul National University, Seoul, Korea
Email: hyoungjuji@islab.snu.ac.kr and bshim@snu.ac.kr
†Dept. of Electrical Engineering, Hanyang University, Seoul, Korea
Email: junwchoi@hanyang.ac.kr

Abstract—In this paper, we propose a novel pilot beamforming and CSI acquisition strategy for machine-type-communication systems to achieve reduction in pilot overhead and improve the channel estimation performance. Key idea of the proposed beamforming technique is to sparsify the time-domain channel vector using the beamforming. To be specific, using the deliberately designed beamforming weight, we minimize non-zero taps and thus sparsify the (beamformed) channel vector. Due to this, only a few samples are required to perform the channel estimation of beamformed pilots so that whole CSI can be acquired with partial pilot symbols in time and frequency. From the numerical evaluations, we show that the proposed scheme outperforms conventional channel estimation schemes, and achieves $N_T$-fold reduction in the pilot overhead. Also, the proposed scheme can be applied for beamforming weight of energy harvesting.

I. INTRODUCTION

Massive machine-type-communication (mMTC), the new use case of 5G communications in which large number of machines, sensors, and vehicles exchange information to improve the quality of human’s life, has received much attention as a baseline for internet of things (IoT) era [1]. To support massive number of machine devices, wireless systems should support various requirements including much higher throughput, lower latency, better reliability, and improved energy efficiency.

One of the main challenges in the design of mMTC systems is that machine devices need to be operated with stringent hardware constraint, power budget, and limited resource constraint. Transceiver in the IoT device consists of the narrowband RF chain, very small number of antenna (at most two), small sized memory, and low-power signal processing unit. Further, machine devices should often enter into the sleep mode or energy harvesting mode to save the energy so that it cannot be possible to feed back the CSI of whole system bandwidth using samples obtained from narrowband measurements. Due to these reasons, systems cannot enjoy the benefit of frequency-selective scheduling, and hence the capability to support massive devices can be affected.

In estimating the channel with limited pilot overhead, compressed sensing (CS)-based pilot transmission and channel estimation techniques have been used popularly [2–5]. While the CS-based channel estimation technique can achieve substantial reduction in the pilot overhead, the overhead is still substantial in the multi-antenna scenarios since the overhead is proportional to the number of transmit antennas at basestation. In fact, since the performance of CS algorithms depend heavily on the sparsity of channel vector in each antenna [3–5], the pilot overhead of the multi-antenna systems increases linearly with the number of antennas $N_T$.

The primary purpose of this paper is to propose a novel pilot beamforming technique for mMTC systems. Proposed method achieves reduction in the pilot overhead and improves the channel estimation performance. Key feature of the proposed pilot beamforming technique is to enforce the sparsity of the time-domain channel vector. To be specific, using the deliberately designed antenna-domain beamforming, we sparsify the time-domain channel vector. From the numerical evaluations, we show that the proposed scheme outperforms conventional channel estimation schemes [3, 6], and also achieves $N_T$-fold reduction in the pilot overhead when the number of the transmit antennas is $N_T$.

II. SYSTEM MODEL

A. Basic System Model

We consider the downlink transmission of mMTC communications with $N_T$ antennas at the basestation and a single antenna at the mobile station. In the OFDM-based system, pilot symbols inserted in the time-frequency grid are used for the channel estimation, data demodulation and channel state information (CSI) feedback. Fig. 1(a) illustrates the resource block (RB) structure of MIMO-OFDM systems. Each box represents a resource element (RE) where a symbol is assigned. In particular, the colored box indicates the RE correspond to the pilot symbols.

Let $\mathbf{y}_n \in \mathbb{C}^{N_P \times 1}$ be the received pilot vector in the frequency domain for the $n$th time-symbol and $i$th antenna (see Fig. 1(b)), then $\mathbf{y}_n$ is expressed as

$$\mathbf{y}_n = \text{diag} \left( \mathbf{p}_n^i \right) \Phi_n^i \mathbf{g}_n^i + \mathbf{z}_n^i,$$

where $\mathbf{p}_n^i \in \mathbb{C}^{N_F \times 1}$ is the pilot symbol vector, $\Phi_n^i \in \mathbb{R}^{N_T \times N_P}$ is the selection matrix containing only one element being one in each row and rest being zero. For example, if the 1st and 3rd subcarriers are used for pilot, then $\Phi_n^i = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \end{bmatrix}$. Also, $\mathbf{z}_n^i \in \mathbb{C}^{N_P \times 1}$ is the additive white Gaussian noise ($\mathbf{z}_n^i \sim \mathcal{CN}(0, \sigma_0^2 \mathbf{I}_{N_P})$) and $\mathbf{g}_n^i \in \mathbb{C}^{N_T \times 1}$ is the frequency-domain channel vector. The relationship between
The angle of departure (AoD) is expressed as the linear array or 2D uniform planner array) and the receiver has a deviation of angular spread for \( h \) and \( n \). The time-domain and frequency-domain channel vectors \( g \) and \( h \) are expressed as

\[
g^i = F h_n^i, \tag{2}
\]

where \( F \) is the DFT matrix.

In characterizing the channel vector \( h \), we assume that there are \( K \) clusters and also \( N_{sp} \) sub-paths for each cluster. When the basestation antennas have a linear array structure (e.g., 1D linear array or 2D uniform planner array) and the receiver has a single antenna, the channel for each antenna can be expressed as the angle of departure (AoD). A time-varying channel tap \( h^i_m \) of the \( i \)th antenna and delay bin \( m \) is expressed as

\[
h^i_m = \frac{1}{N_F} \sum_{l=1}^{N_{sp}} \sqrt{h^i_m e^{j\phi_m} e^{j\theta_m^l \sin \theta_m^l}}, \tag{3}
\]

where \( \kappa = \frac{2\pi}{\lambda} \) is the wavenumber, \( \phi_m \) is the random phase of \( m \)th element, and \( \theta_m^l = \theta_m + \Delta \theta \) is AoD of sub-paths in \( m \)th cluster where \( \Delta \theta \sim \mathcal{N}(0, \sigma_m^2 \), and \( \sigma_m \) is the standard deviation of angular spread for \( m = 1, ..., K \) [7], [8].

Since the distance between antenna elements at the basestation is much smaller than the signal transmission distance in typical multi-antenna geometry, one can assume that channels associated with the transmit-receive antenna pairs share the common support (i.e., \( \text{supp}(h^i) = \text{supp}(h^j) \)) [9].

### B. Beamformed Pilot Transmission

When the number of antennas is large, orthogonal pilot transmission is not so desirable due to the pilot overhead. One option to reduce the pilot overhead is to transmit multiple beamformed pilots [10]. Basic idea of this approach is to transmit the pilot signal after applying the predefined beam pattern. In doing so, multiple beamformed pilots (beams) having different beam directions can be transmitted simultaneously [11], [12]. Since the effective dimension of the beamformed channel vector can be reduced, pilot and feedback overhead can also be reduced. In fact, key idea of the proposed scheme is to design the beamforming weight to make the beamformed channel vector \( h \) sparse.

### III. Sparsification of Pilot Beamforming

In the proposed technique, pilot signal is transmitted after beamforming (see Fig. 2). By the beamforming at the antenna domain, nonzero entries of a time-domain beamformed channel vector is minimized. In this work, we assume that the nonzero tap information is known at the basestation either by the channel feedback or the channel estimation based on the channel reciprocity.

#### A. Time-domain System Model without Pilot Beamforming

Let \( h^i_n = [h^i_n,1, ..., h^i_n, N_{cir}]^T \) be the time-domain CIR vector of \( i \)th antenna. The pilot observation \( y^i_n \) for the conventional wireless systems is

\[
y^i_n = \Phi_n^i g^i_n + z^i_n
\]

\[
y^i_n = \Phi_n^i F \Phi_n^i F^H h^i_n + z^i_n \tag{4}
\]

where \( \Phi \in \mathbb{R}^{N_F \times N_T} \) is the selection matrix containing only one element being one in each column and rest being zero. Let \( U^i_n = \Phi_n^i F \Phi_n^i F^H \), then we have

\[
y^i_n = U^i_n h^i_n + z^i_n \tag{5}
\]

In many wireless environments, the number of nonzero taps \( K \) is much smaller than the length of the channel vector \( N_{cir} \) (i.e., \( K \ll N_{cir} \)). Since \( h^i_n \) is a sparse vector, one can estimate the channel efficiently using the CS technique [2]. In view of multiple antenna systems, the drawback of this approach is that the channel estimation is performed per antenna so that the pilot overhead increase linearly with the number of transmit antenna \( N_T \).

---

1If the first two taps are nonzero elements, then
\[
\Phi = \begin{bmatrix}
1 & 0 & 0 & 0 & \cdots \\
0 & 1 & 0 & 0 & \cdots \\
0 & 0 & 1 & 0 & \cdots \\
0 & 0 & 0 & 1 & \cdots \\
\end{bmatrix}.
\]
B. Pilot Beamforming

When the beamforming weight \( \mathbf{w}_n(k) = [w^1_n(k) \ldots w^{N_T}_n(k)]^T \) is applied, the beamformed (scalar) channel becomes

\[
\tilde{g}_n(k) = \mathbf{w}_n^T(k) \begin{bmatrix} g^1_n(k) \\ \vdots \\ g^{N_T}_n(k) \end{bmatrix}.
\]

Let \( \tilde{\mathbf{h}}_n = [\tilde{h}^1_{n,1} \ldots \tilde{h}^1_{n,N_{crr}}]^T \) be the time-domain CIR vector after the beamforming. Then the received vector \( \tilde{\mathbf{y}}_n \in \mathbb{C}^{N_F \times 1} \) after aggregating all subcarriers can be expressed as

\[
\tilde{\mathbf{y}}_n = \text{diag}(\tilde{\mathbf{p}}_n) \tilde{\Phi}_n \begin{bmatrix} \tilde{g}_n(1) \\ \vdots \\ \tilde{g}_n(N_F) \end{bmatrix} + \tilde{\mathbf{z}}_n
= \text{diag}(\tilde{\mathbf{p}}_n) \tilde{\Phi}_n \text{diag}(\mathbf{W}_n \mathbf{G}_n) + \tilde{\mathbf{z}}_n
= \text{diag}(\tilde{\mathbf{p}}_n) \tilde{\Phi}_n \begin{bmatrix} \tilde{\mathbf{h}}^i_n \\ 0 \end{bmatrix}_{N_F \times N_{crr}} + \tilde{\mathbf{z}}_n,
\]

where \( \tilde{\mathbf{p}}_n \in \mathbb{C}^{N_F \times 1} \) is the pilot symbol vector, \( \tilde{\Phi}_n \in \mathbb{R}^{N_{crr} \times N_{crr}} \) is the selection matrix, \( \tilde{\mathbf{z}}_n \in \mathbb{C}^{N_F \times 1} \) is the additive white Gaussian noise (AWGN) \( \mathbb{C} \mathcal{N}(0, \sigma^2_n \mathbf{I}_{N_F}) \), \( \mathbf{G}_n = [g^1_n \ldots g^{N_T}_n]^T \in \mathbb{C}^{N_{crr} \times N_F} \) is the matrix consisting of frequency-domain channel vectors of \( N_T \) antennas, and \( \mathbf{W}_n = [\mathbf{w}_n(1) \ldots \mathbf{w}_n(N_F)]^T \in \mathbb{C}^{N_F \times N_T} \) is the matrix constructed by stacking beamforming vectors of all subcarriers.

The beamforming weight matrix \( \mathbf{W}_n \) is designed to minimize the nonzero elements (cardinality) of beamformed channel vector \( \tilde{\mathbf{h}}_n \). That is,

\[
\mathbf{W}_n = \arg \min_{\mathbf{W}_n} \| \mathbf{h}_n \|_0
= \arg \min_{\mathbf{W}_n} \left\| \frac{1}{N_F} \mathbf{F}^* \text{diag}(\mathbf{W}_n \mathbf{G}_n) \right\|_0.
\]

Due to the inclusion of \( l_0 \)-norm, the problem seems to be difficult to solve. However, using the deliberately designed beamforming, we can minimize nonzero taps of the time-domain channel vector. For example, if the support (index of nonzero elements) of \( \mathbf{h}^i_n \) is \( \Gamma = \{n_1, n_2, n_3\} \) and the number of antennas \( N_T \) is 4, then the frequency-domain channel for the \( i \)-th antenna and \( k \)-th subcarrier is

\[
g^i_n(k) = \sum_{n=1}^{N_F} h^i_{n,n} e^{\frac{2\pi kn}{N_F}} = [h^i_{n,n_1} h^i_{n,n_2} h^i_{n,n_3}] [e^{\frac{2\pi kn_1}{N_F}} e^{\frac{2\pi kn_2}{N_F}} e^{\frac{2\pi kn_3}{N_F}}]^T.
\]

As mentioned in Section II.A, since the channel gain of the \( i \)-th antenna can be expressed in terms of \( h^1 \) (i.e., \( h^1 e^{j(\theta - \theta_n)} \)), we have

\[
\mathbf{g}^i_n(k) = \begin{bmatrix} g^1_n(k) \\ g^2_n(k) \\ g^3_n(k) \\ g^4_n(k) \end{bmatrix} = \begin{bmatrix} 1 \\ e^{j\theta_1} \\ e^{j\theta_2} \\ e^{j\theta_3} \end{bmatrix} = \begin{bmatrix} h^1_{n,n_1} e^{\frac{2\pi kn_1}{N_F}} \\ h^1_{n,n_2} e^{\frac{2\pi kn_2}{N_F}} \\ h^1_{n,n_3} e^{\frac{2\pi kn_3}{N_F}} \end{bmatrix} = \mathbf{h}_n e^{j\frac{2\pi k}{N_F}}.
\]

\[
\mathbf{g}^i_n(k) = \Omega(k) \mathbf{h}_n.
\]

Let \( \mathbf{g}'_n(k) \) be the channel vector after the beamforming using \( \mathbf{w}_n^i(k) \), the beamformed channel \( \tilde{g}'(k) \) is given by

\[
\tilde{g}'(k) = \mathbf{w}^T_n(k) \mathbf{g}'_n(k) = \mathbf{w}^T_n(k) \Omega(k) \mathbf{h}_n.
\]
In order to generate a sparse beamformed channel such that all taps of $\tilde{h}_i^n$ are zero except for $n_1$-th position, we need to set the beamforming vector as
\[
\mathbf{w}_n^T(k) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \Omega(k)^\dagger \approx \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} (\Omega^T(k)\Omega(k))^{-1}\Omega^T(k).
\]
(13)

Since the dimension of matrix $\Omega(k)$ is $N_T \times K$ and $\Omega(k)$ is a full rank matrix, we can compute the pseudo-inverse as long as $N_T \geq N$. Note that other than mMTC scenarios, this condition is also well suited for millimeter wave communication scenarios and energy harvesting networks. After the beamforming, we obtain the beamformed frequency-domain (scalar) channel $\tilde{g}_n(k)$ as
\[
\tilde{g}_n(k) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} (\Omega^T(k)\Omega(k))^{-1}\Omega^T(k)\Omega(k)\tilde{h}_n
\]
\[
= \tilde{h}_{n,n_1}e^{j\frac{2\pi n_1 k}{N_T}}, \quad K \leq N_T.
\]
(14)

By converting the beamformed frequency-domain channel vector $\tilde{g}_n$ into the time-domain, we obtain sparse channel vector in time-domain $\hat{h}_n = \frac{1}{N_T} F^* \tilde{g}_n = [0 \ldots 0 \tilde{h}_{n,n_1} \ldots 0]^T$. Without loss of generality, to select $K$ taps out of $K$ taps, we can use binary vector $e$ as
\[
\mathbf{w}_n^T(k) = e^T \Omega(k)^\dagger.
\]
(15)

Now, by transmitting multiple beams simultaneously and then choosing the beam with the best quality, one can further improve the performance. When the number of simultaneously transmitted beam is $J$, the aggregated time-domain channel vectors of beams can be expressed as
\[
[\hat{h}_n^1 \ldots \hat{h}_n^J] = \frac{1}{N_F} F^* \left[ \text{diag}(\mathbf{W}_n^1 \mathbf{G}_n) \ldots \text{diag}(\mathbf{W}_n^J \mathbf{G}_n) \right]
\]
\[
= \frac{1}{N_F} F^* \left[ \begin{array}{c}
\mathbf{w}_n^1(1) \ldots \mathbf{w}_n^J(1) \\
\vdots \\
\mathbf{w}_n^1(N_F) \ldots \mathbf{w}_n^J(N_F) \end{array} \right] \Omega(1) \hat{h}_n
\]
\[
= \frac{1}{N_F} F^* \left[ \begin{array}{c}
\mathbf{w}_n^1(N_F) \ldots \mathbf{w}_n^J(N_F) \end{array} \right] \Omega(1) \hat{h}_n
\]
(16)

In each subcarrier $k$, the precoding weights are expressed as
\[
[\mathbf{w}_n^1(k) \ldots \mathbf{w}_n^J(k)] = E\Omega(k)^\dagger \quad \text{for} \quad K \leq N_T.
\]
(17)

where $E = [e^1 \ldots e^J]^T$. For example, by setting
\[
\text{supp } \{e^i\} \neq \text{supp } \{e^j\} \quad \text{for } \ i, j = 1, \ldots, J,
\]
(18)
each of beamformed pilot signals is placed in a separate tap of the time-domain channel vector. This setting is useful especially when an MTC device performs the channel estimation for multiple beamformed pilots.

IV. SIMULATIONS AND DISCUSSION

We consider OFDM-based mMTC systems where $B_s = 20$ MHz, subcarrier spacing of 15 KHz, and DFT size of $N_F = 2048$. mMTC device uses 12 subcarriers ($B_u = 150$ KHz). The maximum delay spread of the multipath channel is assumed to be 0.467 $\mu$s, which yields $N_{cir} = 144$. Duration of symbol is 7.2 $\mu$s and the time interval between adjacent pilot symbols is 5 ms. We assume that 5 taps in the channel vector are dominant with uniform energy. One subcarrier is randomly chosen for the pilot purpose among the $B_u$ bandwidth. In the proposed beamforming, 5 beams are used for the pilot transmission.

We first measure the MSE performances of channel estimation with and without pilot beamforming (see Fig. 3). In case of the pilot beamforming, multiple beams are used and time-domain channel vectors $\hat{h}_n$ are jointly estimated from 5 beamformed pilots. Without the pilot beamforming, pilots are transmitted per antenna and time-domain channel vectors $\hat{h}_n^i$ are estimated for each antenna $i$. We observe that the proposed scheme outperforms the conventional schemes and achieves more than 20dB gain in high SNR regime. The benefit of sparsification can be better understood by observing the normalized throughput. In Fig. 4, we plot the normalized throughput of the proposed scheme. As a performance metric, we define the normalized throughput $\eta$ as
\[
\eta \text{ (bps/Hz)} = \frac{N_{bit}}{(N_t - N_p)T_d B_s},
\]
(19)
where $N_{bit}$ is the number of successive bit during $T_d$ duration, $N_t$ is the number of REs used for transmission, and $N_p$ is the number of REs used for pilot transmission. In the pilot beamforming, one subcarrier in 12 subcarriers ($B_u$ bandwidth) is used for pilot transmission. For example, to use $J$ beams, this requires $JK$ pilot subcarriers. In the conventional CS scheme, $N_tK$ subcarriers are used for pilot transmission. For the data transmission, we use the best beam feed back from the MTC device [10]. Simulation results demonstrate that the proposed scheme outperforms the conventional scheme without beamforming even when the pilot overhead is accounted for.
V. CONCLUSION

In this paper, we proposed an efficient pilot beamforming scheme for the mMTC systems. The key feature of the proposed pilot beamforming scheme is to minimize cardinality of beamformed channel vectors using the antenna-domain beamforming. The received channel vector of beamformed pilots can be jointly estimated by the sparse recovery algorithm. From simulations, we observed that the proposed scheme achieves significant reduction in the pilot overhead for realistic mMTC scenarios. Furthermore, the proposed scheme can be applied mMTC who harvests energy from the beamformed channel.

REFERENCES