HUMAN AND MACHINE TYPE COMMUNICATIONS CAN COEXIST IN UPLINK MASSIVE MIMO SYSTEMS

Kamil Senel, Emil Björnson and Erik G. Larsson

Department of Electrical Engineering,
Linköping University, Linköping, Sweden
Email: {kamil.senel, emil.bjornson, erik.g.larsson}@liu.se

ABSTRACT

Future cellular networks are expected to support new communication paradigms such as machine-type communication (MTC) services along with human characteristics (HTC) services. This requires base stations to serve a large number of devices in relatively short channel coherence intervals which renders allocation of orthogonal pilot sequence per-device approaches impractical. Furthermore, the stringent power constraints, place-and-play type connectivity and various data rate requirements of MTC devices make it impossible for the traditional cellular architecture to accommodate MTC and HTC services together. Massive multiple-input-multiple-output (MaMIMO) technology has the potential to allow the coexistence of HTC and MTC services, thanks to its inherent spatial multiplexing properties and low transmission power requirements. In this work, we investigate the performance of a single cell under a shared physical channel assumption for MTC and HTC services and propose a novel scheme for sharing the time-frequency resources. The analysis reveals that MaMIMO can significantly enhance the performance of such a setup and allow the inclusion of MTC services into the cellular networks without requiring additional resources.

Index Terms— MIMO, Machine Type Communication.

1. INTRODUCTION

One of the key technologies of 5G future networks is the machine-type communications (MTC), which is projected to provide wireless connectivity to tens of billions of devices as a result of smart cities, factories, vehicles, and even common objects with sensing and communicating capabilities [1]. A potential solution for accommodating the emerging traffic is utilizing the already existing infrastructure of cellular networks. The standardization of techniques for MTC over cellular networks is already being considered [2]. However, the existing cellular network architectures, which are optimized to handle human-type communications (HTC), must be modified in order to handle HTC along with MTC, which requires consideration of a diverse communication characteristics [3].

There are crucial problems to solve before the successful integration of MTC services into the existing cellular networks. In particular, 5G networks will have to support a large number of devices with low-complexity constraints and provide various data rates ranging up to multiple gigabits per second with reliability for services with stringent latency constraints such as health-care, security, and automotive applications [4].

The performance of cellular networks in a setup where HTC and MTC services coexist has been considered in [1][5]. WiFi-based networks constitute a competitive option to cellular networks and the integration of MTC services into the existing WiFi-based networks has been investigated [6]. A potential alternative is to utilize multihop short-range transmission technologies [7]. However, initial experiments reveal the limitation of short-range technologies for MTC applications and emphasized the requirement of a plug-and-play type of connectivity without centralized planning which can be satisfied by long-range technologies [8].

A key technology of 5G future cellular networks is MaMIMO in which the BSs are equipped with a large number of antennas, which gives them the ability to spatially multiplex multiple users [9]. The advantages of MaMIMO technology have been shown to enhance the performance of cellular networks in terms of spectral efficiency and device detection in MTC setups [10][11][12]. However, to the best of authors’ knowledge, this is the first work which considers the coexistence of HTC and MTC devices under a MaMIMO setup and analyze their joint spectral efficiency.

In this work, we investigate the performance, in terms of spectral efficiency, of a MaMIMO network that concurrently serves devices that utilize HTC and MTC. Different schemes for allocating time-frequency resources between MTC and HTC devices are investigated. A novel resource allocation scheme is proposed and compared with the orthogonal and non-orthogonal resource allocation schemes. In particular, we answer the following questions:

- How will the existing cellular networks be affected by the dense MTC deployments?
- Can the challenges to accommodate MTC services over cellular networks be handled by the MaMIMO technology?
- Does the MaMIMO technology enable the use of non-orthogonal resources for MTC and HTC services thanks to its inherent utilization of spatial multiplexing?

2. SYSTEM SETUP

We consider the uplink of a single-cell MaMIMO system where a BS with M antennas is serving K devices. Among these devices K_m of them, referred to as machines, require machine-type communication and the remaining K_h = K - K_m devices, referred to as humans, generate human-type traffic. Without loss of generality, assume that the indexes k \in \mathcal{X}_h, where \mathcal{X}_h = \{1, \ldots, K_h\} are utilized for humans and k \in \mathcal{X}_m = \{K_h + 1, \ldots, K\} denotes machines.

The time-frequency resources are divided into coherence intervals, such that each channel is constant and frequency-flat in each interval. Each coherence interval (CI) has length N (in samples) and a fraction of these samples are reserved for training whereas the remaining ones are utilized for uplink data transmission.
Humans
Machines
Humans
Machines
Pilots
Data
Pilots
Data
Pilots
Data
Pilots
Data

Non-line-of-sight communication is assumed and the small-scale fading is modeled as Rayleigh fading. Let \( \mathbf{g}_k \in \mathbb{C}^{M \times 1} \) denote the channel between device \( k \) and the BS, then

\[
\mathbf{g}_k = \sqrt{\beta_k} \mathbf{h}_k, \quad \forall k = 1, \ldots, K,
\]

where \( \beta_k \) is the large-scale fading (same for all antennas) and \( \mathbf{h}_k \) is the small-scale fading. Each element of \( \mathbf{h}_k \) is modeled as i.i.d. \( \mathcal{CN}(0, 1) \). The BS is assumed to know the large-scale fading coefficients. However, the small-scale fading coefficients are to be estimated in each CI. Moreover, \( \mathbf{h}_k \) changes independently between CIs.

2.1. CI Allocation Schemes and Pilot Sequences

We consider three training and data transmission schemes.

- **Scheme 1:** Humans and machines utilize different CIs.
- **Scheme 2:** All devices use the same training interval and data transmission interval in every CI.
- **Scheme 3:** Machines are not allowed to transmit during the training period of humans, which reduces the human’s pilot length. After the training of humans, machines transmit their pilot sequences followed by data transmission.

Fig. 1 illustrates the CI structures for the three schemes. Scheme 1 is an orthogonal allocation scheme in the sense that it allocates different CIs to humans and machines. Scheme 2 and 3 are non-orthogonal schemes where both machines and humans utilize the same CIs. In Scheme 3, we propose a novel approach by utilizing the training period of machines for data transmission of humans.

In MaMIMO setups, all \( K \) devices concurrently transmit their pilot sequences during uplink training. The channels are estimated during uplink training and utilized during uplink data transmission. Traditionally, the pilot sequences are assumed to be orthogonal for users within a cell. These assumptions on pilot sequences are not realistic for massive MTC since it requires many pilots and cumbersome access procedures for pilot assignment. Moreover, constraints on uplink power budget and excessive overhead signaling compel the use of non-orthogonal pilots for MTC.

It is assumed that the number of machines is larger than the number of humans \( K_m > K_h \) and it is not feasible to assign orthogonal pilots to machines. However, since the humans require higher data rates and are smaller in numbers compared to machines, we assume it is possible to allocate orthogonal pilot sequences to the \( K_h \) humans. Let \( \sqrt{N_p} \varphi_k \in \mathbb{C}^{N_p \times 1} \) denote the \( N_p \)-length pilot signal for the \( k \)th device with \( \| \varphi_k \|^2 = 1 \) and assume \( k \in \{1, \ldots, K_h\} \), then

\[
\varphi_k^H \varphi_i = 0, \quad \forall k \in \{1, \ldots, K_h\}, \forall i \in \{1, \ldots, K\}.
\]

For the MTC applications such as surveillance, machines also require high data-rates. In these cases the machines with high data rate requirements may be treated as humans.

The pilot sequence length of humans, \( N_p^h \), must be equal to or greater than the number of humans \( K_h \) in order to allocate orthogonal pilots. On the other hand, machines are assumed to share \( N_p^m \) orthogonal pilots. Each device picks one of the \( N_p^m \) pilot sequences randomly in each coherence interval. Let \( \varphi_k \) be a pilot sequence assigned to a machine, then

\[
\varphi_k^H \varphi_i = \begin{cases} 
0, & \text{if } i \in \mathcal{K}_h, \\
0, & \text{with Pr. } 1 - \frac{1}{N_p^m}, \text{ if } i \in \mathcal{K}_m, \\
1, & \text{with Pr. } \frac{1}{N_p^m}, \text{ if } i \in \mathcal{K}_m.
\end{cases}
\]

Although this pilot allocation avoids the need of a coordination process between the BS and the active devices in each CI, it introduces pilot contamination between machines. Alternatively, orthogonal pilots can be assigned to active machines in each CI. However, this requires resources such as additional symbols and transmit power for coordination which are not desirable for MTC with latency and power constraints.

3. ACHIEVABLE RATE ANALYSIS

In this section, the achievable rate of the three schemes illustrated in Fig. 1 is investigated. Each scheme has an uplink training phase followed by data transmission.

3.1. Analysis of Scheme 1

In Scheme 1 (SC-1), humans and machines utilize different CIs and there is no interference between them. Let \( \mathcal{K} \) be the set of active devices in a given CI. During the uplink training, active devices \( k \in \mathcal{K} \) (in SC-1, either humans or machines are active, i.e., \( \mathcal{K} = \mathcal{K}_h \) or \( \mathcal{K} = \mathcal{K}_m \)) concurrently transmit their pilot sequences and the composite received signal at the BS is

\[
\mathbf{Y} = \sqrt{N_p} \sum_{k' \in \mathcal{K}} \sqrt{q_{k'}} \mathbf{g}_{k'} \varphi_{k'}^H + \mathbf{Z}
\]

where \( \mathbf{Z} \in \mathbb{C}^{M \times N_p} \) is the noise matrix with i.i.d. \( \mathcal{CN}(0, \sigma^2) \) elements and \( q_{k'} \) denotes the transmission power of pilot symbols for user \( k' \).

In order to estimate the channel of user \( k \), the BS performs a de-spreading operation on the received signal:

\[
\mathbf{y}_k = \mathbf{Y} \varphi_k = \sqrt{N_p} q_k \mathbf{g}_k + \sqrt{N_p} \sum_{k' \in \mathcal{K} \setminus \{k\}} \sqrt{q_{k'}} \mathbf{g}_{k'} \varphi_{k'}^H + \mathbf{z'}
\]

where \( \mathbf{z'} = \mathbf{Z} \varphi_k \) has i.i.d. \( \mathcal{CN}(0, \sigma^2) \) elements since \( \| \varphi_k \|^2 = 1 \). Then, the BS utilizes the linear minimum mean-square error (LMMSE) estimator to acquire \( \hat{\mathbf{g}}_k \), the estimate of the channel between the BS and user \( k \), as follows:

\[
\hat{\mathbf{g}}_k = \frac{\sqrt{N_p} q_k \beta_k}{\sum_{k' \in \mathcal{K}} q_{k'} \beta_{k'} \| \varphi_{k'}^H \varphi_k \|^2 + \sigma^2} \mathbf{y}_k,
\]
(which is the true MMSE estimate for Scheme-1 and Scheme-2). \( \hat{g}_k \) has \( M \) i.i.d. elements and the mean-square of the \( m \)th component is

\[
\gamma_k = \mathbb{E}_g \left[ |\hat{g}_k|_m^2 \right] = \frac{N_p q_k \beta_k^2}{N_p \sum_{k' \in \mathcal{K}_k} q_k \beta_k |\varphi_k^H \varphi_{k'}|^2 + \sigma^2},
\]

where \( \mathbb{E}_g \) denotes the expectation with respect to \( g \). Let \( s_k \) denote the unit power symbol to be conveyed by device \( k \). Then, device \( k \) transmits \( s_k = \sqrt{P_k} x_k \), where \( P_k \) is the data transmit power of device \( k \). In order to detect the data symbols of the \( k \)th device, the BS employs the maximum ratio combining (MRC) with the combining vector

\[
\hat{v}_k = \frac{1}{\gamma_k} \sqrt{M} \hat{g}_k
\]

to compute the inner product with the received signal, \( y \) as

\[
\hat{v}_k^H y = \hat{v}_k^H \sqrt{P_k} x_k + \sum_{k' \in \mathcal{K}_k \setminus \{k\}} \hat{v}_k^H \sqrt{P_k} x_{k'} + \hat{v}_k^H z.
\]

Based on [9], the achievable rate of device \( k \) is given by

\[
R_k = \alpha_k \left( \frac{N_d}{N} \right) \log_2 (1 + \Gamma_k),
\]

where

\[
N_d = \begin{cases} N - N_k^p, & \text{if } k \in \mathcal{K}_h, \\ N - N_m^p, & \text{if } k \in \mathcal{K}_m. \end{cases}
\]

Here, \( N_k^p \) and \( N_m^p \) are the pilot lengths for humans and machines, respectively. \( \alpha_k \in [0, 1] \) represents the fraction of CIs assigned to humans/machines. The effective SINR term in [10] is [15]

\[
\Gamma_k = \begin{cases} \frac{M p_k}{N_k'}, & \text{if } k \in \mathcal{K}_h, \\ \frac{1}{N_k'} \left( \sum_{k' \in \mathcal{K}_h} p_{k'} \beta_{k'} + \sigma^2 \right), & \text{if } k \in \mathcal{K}_m, \end{cases}
\]

where \( \mathcal{K}_m = \mathcal{K}_k \setminus \{k\} \) and

\[
\tilde{\gamma}_k = \mathbb{E}_g \left[ \frac{1}{\gamma_k} \right]^{-1} = \frac{N_m q_k \beta_k^2}{N_p q_k \beta_k + \sum_{k' \in \mathcal{K}_h} q_k \beta_k |\varphi_k^H \varphi_{k'}|^2 + \sigma^2}.
\]

3.2. Analysis of Scheme 2

In Scheme 2, each device uses \( N_p \) symbols for training and \( N - N_p \) symbols for data. To find the corresponding rate of device \( k \), we utilize the bounding techniques given in [2] and state the following:

**Lemma 1** The achievable rate of device \( k \) under Scheme 2 is

\[
R_k = \left( \frac{N - N_p}{N} \right) \log_2 (1 + \Gamma_k),
\]

where \( \Gamma_k \) is the effective SINR for device \( k \) and is defined by

\[
\Gamma_k = \begin{cases} \frac{M p_k}{N_k'}, & \text{if } k \in \mathcal{K}_h, \\ \frac{1}{N_k'} \left( \sum_{k' \in \mathcal{K}_h} p_{k'} \beta_{k'} + \sigma^2 \right), & \text{if } k \in \mathcal{K}_m, \end{cases}
\]

3.3. Analysis of Scheme 3

In Scheme 3, the machines are silent during the training of humans and send their pilot sequences while humans are transmitting data. The LMMSE channel estimate for machines is

\[
\hat{g}_k = \sqrt{\frac{N_m q_k \beta_k}{N_p \sum_{k' \in \mathcal{K}_m} q_k \beta_k |\varphi_k^H \varphi_{k'}|^2 + \sigma^2}} y_k.
\]

The effective SINRs given by [15] reveals that as long as orthogonal pilots are assigned to humans, the integration of machines to the existing network does not create coherent interference, i.e., interference that scales with the number of antennas. Hence, as \( M \) grows, the additional interference originating from machines vanishes. However, this is not the case for machines as they suffer coherent interference due to the use of non-orthogonal pilots.

3.4. Asymptotic Analysis

In order to gain further insights, the asymptotic limits of the rate expressions as \( M \to \infty \) are investigated in this section. Note that as \( M \to \infty \), the rate of humans, \( R_k \to \infty \), \( \forall k \in \mathcal{K}_h \) in all of the schemes considered. This is to be expected as humans suffer no coherent interference thanks to the orthogonal pilots allocated for them. For machines, the asymptotic limits are summarized as follows.
Corollary 1 The achievable SINR for device $k \in \mathcal{K}_m$ as $M \to \infty$ is given by

$$\Gamma_k = \left\{ \begin{array}{ll}
\frac{p_k}{\sum_{k' \in \mathcal{K}_m} p_{k'} h_{k'k}^2}, & \text{for SC-1 and SC-2}, \\
\frac{1}{N_p K_m} \left( \frac{1}{\sum_{k' \in \mathcal{K}_m} h_{k'k}^2} + \frac{1}{\sum_{k' \in \mathcal{K}_m} h_{k'k}^2} \right), & \text{for SC-3}. 
\end{array} \right. \tag{21}$$

The proof follows from taking the limit in (12), (15) and (20). The asymptotic analysis shows that SC-1 and SC-2 are equivalent in terms of asymptotic SINR whereas in SC-3 machines suffer additional coherent interference originating from humans. In both cases, the effective SINR increases with the pilot length of machines.

4. NUMERICAL RESULTS

In this section, numerical results are presented which validate the analytical results in Section 3 along with the simulations that compare the schemes introduced in Section 2. The simulation setup consists of a cell with radius 250 m serving $K_h = 5$ humans and $K_m = 15$ machines which are uniformly distributed with a guard interval of 20 m. The path loss at distance $d$ (km) is given by $130 + 37.6 \log_{10}(d)$ in dB. The power spectral density of the noise is $-174$ dBm/Hz and the maximum transmit power for each device is 30 dBm.

Fig. 2 depicts the rate regions for the schemes described in Section 3. Here, $R_h$ and $R_m$ denote the max-min rate for humans and machines respectively. The rate curves are obtained by maximizing the minimum rate with respect to the transmit powers and the machine pilot length $N_p^m$. The pilot length of humans is fixed at $N_p^h = K_h$. The rate regions are obtained for two different CI lengths, which illustrates that Scheme-1 performs the best when there is only a single type of active device, i.e., either only humans or machines, in a given CI. However, for the cases where machines and humans coexist, allowing transmission from both results in a higher ergodic achievable rate as illustrated by Scheme-2 and Scheme-3. Furthermore, the numerical analysis shows that the novel Scheme-3 outperforms the other schemes under a heterogeneous setup.

The asymptotic limits provided in Corollary 1 and the ergodic achievable rate with respect to the number of BS antennas are depicted in Fig. 3. Note that in each of the schemes $R_h \to \infty$ as $M \to \infty$ and therefore only the ergodic achievable rates of machines are included in the simulations which illustrates that even for relatively low number of antennas the difference between the asymptotic and achieved rate is small. This deviates from the conventional asymptotic results in MaMIMO, which require a very large number of antennas to reach asymptotic limits [15].

The impact of the number of antennas is illustrated in Fig. 3. As $M$ increases the non-orthogonal schemes (SC-2, SC-3) outperform orthogonal scheme (SC-1) due to two important reasons. First, the interference between humans and machines decreases with $M$, effectively converging to the SINR in the orthogonal scheme as $M \to \infty$. Also, the pre-log factor becomes dominant with increasing $M$ as the system starts to operate in bandwidth limited region at higher $M$ values contrary to lower $M$ values in which the system is in power limited region.

5. CONCLUSION

In this work, we consider the problem of accommodating both MTC and HTC in a MaMIMO system and proposed a novel resource allocation scheme along with the analysis of achievable rates for HTC and MTC. The analysis shows that MaMIMO is a key technology for enabling MTC in cellular networks thanks to the large array gain and spatial multiplexing.
6. REFERENCES

[1] Z. Dawy, W. Saad, A. Ghosh, J. G. Andrews, and E. Yaacoub, “Toward massive machine type cellular communications”, IEEE Wireless Communications, vol. 24, no. 1, pp. 120–128, 2017.

[2] 3rd Generation Partnership Project (3GPP), Technical Specification Group GSM/EDGE Radio Access Network, Cellular System Support for Ultra Low Complexity and Low Throughput Internet of Things, Release 13, June 2015.

[3] METIS D6.6, “Final Report on the METIS System Concept and Technology Roadmap”, 2015.

[4] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, et al., “Scenarios for 5G mobile and wireless communications: the vision of the METIS project”, IEEE Communications Magazine, vol. 52, no. 5, pp. 26–35, 2014.

[5] C. S. Bontu, S. Periyalwar, and M. Pecen, “Wireless wide-area networks for internet of things: An air interface protocol for IoT and a simultaneous access channel for uplink IoT communication”, IEEE vehicular technology magazine, vol. 9, no. 1, pp. 54–63, 2014.

[6] G. J. Sutton, R. Liu, and Y. J. Guo, “Harmonising Coexistence of Machine Type Communications with Wi-Fi Data Traffic under Frame-based LBT”, IEEE Transactions on Communications, 2017.

[7] M. Centenaro, L. Vangelista, A. Zanella, and M. Zorzi, “Long-range communications in unlicensed bands: The rising stars in the IoT and smart city scenarios”, IEEE Wireless Communications, vol. 23, no. 5, pp. 60–67, 2016.

[8] A. Biral, Marco Centenaro, Andrea Zanella, Lorenzo Vangelista, and Michele Zorzi. “The challenges of M2M massive access in wireless cellular networks.” Digital Communications and Networks 1, no. 1 (2015): 1-19.

[9] T. L. Marzetta, E. G. Larsson, H. Yang and H. Q. Ngo, “Fundamentals of Massive MIMO”, Cambridge University Press, 2016.

[10] Z. Jiang, B. Han, P. Chen, F. Yang, and Q. Bi, “On Novel Access and Scheduling Schemes for IoT Communications”, Mobile Information Systems, 2016.

[11] E. De Carvalho, E. Björnson, J. H. Sorensen, E. G. Larsson, and P. Popovski, “Random Pilot and Data Access in Massive MIMO for Machine-type Communications”, IEEE Transactions on Wireless Communications, 2017.

[12] K. Senel and E. G. Larsson, “Device Activity and Embedded Information Bit Detection Using AMP in Massive MIMO”, arXiv preprint [arXiv:1709.05161] (2017).

[13] E. De Carvalho, E. Björnson, J. H. Sorensen, P. Popovski, and E. G. Larsson, “Random Access Protocols for Massive MIMO”, IEEE Communications Magazine, vol. 55, no. 5, pp. 216–222, 2017.

[14] R. Ratasuk, J. Tan, and A. Ghosh, “Coverage and capacity analysis for machine type communications in LTE”, in IEEE Vehicular Technology Conference, pp. 1–5, 2012.

[15] D. Verenzuela, E. Björnson, and L. Sanguinetti, “Spectral and Energy Efficiency of Superimposed Pilots in Uplink Massive MIMO”, arXiv preprint [arXiv:1709.07722] (2017).