"POLITEHNICA" UNIVERSITY OF BUCHAREST

PhD SCHOOL ETTI-B

Decision No. 537 from 28.07.2020

DOCTORAL THESIS

-SUMMARY-

Algoritmi și Proceduri Inovative în Sisteme de Comunicații Digitale 5G

Innovative Algorithms and Procedures for 5G Digital Communications Systems

PhD Student: **Eng. Răzvan-Florentin TRIFAN**

DOCTORAL COMMITTEE

President	**Prof. dr. eng. Gheorghe BREZEANU**	from	**"Politehnica" Univ. of Bucharest**
PhD supervisor	**Prof. dr. eng. Constantin PALEOLOGU**	from	**"Politehnica" Univ. of Bucharest**
Reviewer	**Prof. dr. eng. Daniela TĂRNICERIU**	from	**"Gheorghe Asachi" Technical Univ. of Iași**
Reviewer	**Prof. dr. eng. Ioan NICOLAESCU**	from	**"Ferdinand I" Military Technical Acad. of Bucharest**
Reviewer	**Prof. dr. eng. Călin VLĂDEANU**	from	**"Politehnica" Univ. of Bucharest**

Bucharest 2020

Contents

1 Introduction

In telecommunications domain, 5G is the 5^{th} generation of cellular networks technologies. Multiple antennas [Multiple-Input-Multiple-Output (MIMO)] systems are used in 5G networks to increase the capacity and coverage of the radio link.

Digital signal processing is at the core of MIMO systems, used for improving the stability and the capacity of the radio links. Precoding is a digital signal processing technique applied at transmission. It is used mostly in MIMO systems with multiple users (MU-MIMO), where the random location of users does not enable the simultaneous decoding of the multiple data flows received from the transmit antennas.

MIMO techniques are the digital signal processing schemes used for the transmission of multiple data flows through the radio channel, using the same frequency and time resources. The term precoding will be used to describe any of the techniques above, applied to the transmitted signals in MU-MIMO systems.

1.1 Purpose of the thesis

A first objective of the thesis is the identification of simulation methods for the MU-MIMO systems. As such, various scenarios and channels models were proposed to evaluate the precoding algorithms performance. A second objective of the thesis is the study of the performance for the existing precoding schemes in the literature, in spatial and temporal correlated channels. This study led to the identification and the performance optimization of a hybrid precoding scheme. The third and the main objective of this work is the performance enhancement of a hybrid precoding scheme through a reduced complexity method that identifies the groups of users that are spatially separable. This led to the implementation of a statical users grouping algorithm based of their spatial compatibility. For an accurate performance evaluation of the precoding scheme and of the users clustering algorithm, another objective of the thesis is the study of a method for generating realistic traffic distributions, using measurements collected from call traces, geolocated in three dimensions (3D).

1.2 Content of the thesis

The thesis is organized as follows. In Chapter 2, the main concepts related to 5G communication systems are described. Chapter 3 presents the basic theoretical knowledge required for understanding the MIMO domain. Channel modeling, as well as a method for generating a spatial and temporal correlated channel, are introduced. In Chapter 4, the main precoding techniques are summarized and their performance in terms of spectral efficiency and computational complexity, is then evaluated. In Chapter 5, an algorithm for grouping of users based on their spatial correlation is proposed. The performance in terms of computational complexity and spectral efficiency

of the grouping algorithm and of the hybrid precoding scheme are then evaluated. Chapter 6 presents a method for generating a realistic distribution of users in space, which is used in the performance evaluation of the clustering algorithm and of the hybrid precoding scheme from Chapter 5. Chapter 7 presents a method of selection and ordering of users, by defining a fully configurable scheduler, which can be tuned based of the performance of a real base station implementation. Chapter 8 summarizes the main results, the original contributions and the future development perspectives.

2 5G communications systems

The 5G standard developed by 3GPP is called New Radio (NR). Before its final version being published, version 15, the key concepts resulted from the documents of the 3GPP radio group's meetings were published in [1].

Scenarios and technologies used in 5G

The deployment scenarios of the services used in 5G communications systems are classified in three main categories:

- Ultra Reliable and Low Latency Communications (URLLC);
- massive Machine-Type Communications (mMTC);
- enhanced Mobile Broadband (eMBB).

eMBB services are characterized by high bit rate, high user density, high coverage and mobility.

A different set of requirements for 5G systems are used by the stable connections, that require low latency and precise locations for ensuring services for public safety, like controlling and ensuring the safety for the car traffic, or remote medical diagnosis. The enabling of mMTC communications requires the support for a high density of connections, with low and variable bit rate.

Ensuring support for the above scenarios translates into a set of technology requirements.

Millimeter Waves

For reaching a Gbps order of bit rates, new frequency bands in the 30 - 300 GHz range (called millimeter waves or mmWaves) are used, where there is enough unused spectrum. As in these bands the signal attenuation is higher, multiple antenna systems with a high number of antenna elements are required for beam steering and beamforming, in order to compensate for the higher pathloss.

In addition to a higher pathloss, in mmWaves bands, another issue is the interference between Orthogonal Frequency Division Multiplex (OFDM) sub-carriers, which

is caused by a higher phase noise in high frequency bands. The solution chosen consists in using additional reference signals like Tracking Reference Signals (TRS) or increasing the sub-carrier bandwidth, from 15 kHz as in LTE, to 120 or 240 kHz.

Variable numerology

Compared to LTE, NR uses an OFDM modulation with a variable numerology. The distance between the OFDM sub-carriers is no longer fixed to 15 kHz as in LTE, but can reach 240 kHz. The different numerologies can be multiplexed in time and in frequency within the same frame, in order to satisfy various services requirements. This way, URLLC services will be deployed in low and mid bands using a higher numerology, such that the symbols used to be narrower and to reduce the network access time. On the other front, eMBB services will use lower numerologies, having a longer transmission duration.

Multiple antennas

In addition to the use of wider band channels, the bit date can be enhanced through spatially multiplexed data flows, transmitted by the multiple antennas of a base station, over the same time and frequency resources.

To compensate for the higher pathloss of mmWaves through beamforming, in 5G MIMO systems, a hybrid architecture of analog-digital beamforming is proposed. This way the possibility of doing spatial multiplexing is kept, without increasing too much the cost of the antenna system.

In MIMO precoding (as will be presented in Chapter 4 of this thesis), the number of radio chains of a multiple antenna system represents the number of virtual antennas available for spatial multiplexing, whereas the number of antenna elements of a radio chain determines the maximum beamforming gain.

Massive MIMO

Massive MIMO multiple antenna systems were proposed to address the massive capacity increase of the 5G networks. They are MU-MIMO communications systems of type MU-MIMO, with a larger number of transmit antennas than the number of users, much larger than the classical MIMO systems, being capable of doing simultaneously spatial multiplexing and beamforming [2].

Besides the increase of the received power, according to [3], the advantages of using Massive MIMO are:

- Radio link capacity increased by ten times, and the radiated energy efficiency increased by a hundred times;
- Reduced cost of the lower power radio components;
- Drastic reducing of the radio interface latency. Massive MIMO systems are functioning based on the law of large numbers and on beamforming in order to avoid strong fading channels and increased latency;

- Multiple access scheme simplification. With OFDM, each sub-carrier of a Massive MIMO system could exhibit the same channel gain, as such every user can use the entire bandwidth and thus, simplify the signaling of the radio physical channel resources by the network.

The factors that limit the performance of the Massive MIMO systems are:

- Radio chains complexity (made out of 100-200 parallel paths) assuming a fully digital configuration;
- Size of the digital-analog interface (which should be able to process around 100-200 transmit-receive radio paths);
- Information quantity for channel estimation (information without any use for the user);
- High complexity of the digital signal processing;
- Time Division Duplex (TDD) only is used to estimate the Downlink (DL) channel using the estimate of the Uplink (UL) channel;
- Pilot signal contamination.

3 Elements of MIMO theory

This chapter describers the main concepts of the MIMO domain. A method for channel modeling together with a method for creating spatially and temporally correlated channels, are introduced. A way of estimating the spatial separability between users is also presented.

3.1 MIMO channel

The channel matrix of a MIMO antenna system with N_R receive antennas and N_T transmit antennas is written as $\mathbf{H}$, $\mathbf{H} \in C^{N_R \times N_T}$. MU-MIMO architecture is also a MIMO system, but in this case the N_R antennas are distributed to K users.

Assuming $\mathbf{x}$ is the set of symbols being transmitted independently by the N_T antennas, $x_1, x_2, ..., x_{N_T}$, $\mathbf{x} \in \mathbb{C}^{N_T \times 1}$, the signal received by the N_R antennas is $\mathbf{y} \in \mathbb{C}^{N_R \times 1}$, and can be obtained as:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{3.1}$$

where $\mathbf{n} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}_{N_R}, N_0\mathbf{I}_{N_R})$ is the noise received by the N_R antennas, having a power spectral density N_0. It is assumed that it follows a complex Gaussian distribution, circular symmetric[1]with zero mean.

[1]An array $\mathbf{n}$ is circular symmetric if $e^{j\theta}\mathbf{n}$ has the same distribution as $\mathbf{n}$, for every θ.

The ideal channel modeling of type Rayleigh (from a MIMO point of view), assumes a perfect decorrelation of the signals received by the K mobile stations, where all the elements of the matrix $\mathbf{H}$ are complex random variables, Gaussian distributed, $h_{ij} \sim \mathcal{N}_{\mathbb{C}}(0,1)$.

3.2 Channel spatial correlation

In the MU-MIMO context, a random array $\mathbf{h}_k$ is spatially correlated using the Karhunen-Loeve transformation [4]:

$$\mathbf{h}_k = \mathbf{R}_k \widehat{\mathbf{e}}_k = \mathbf{U}_k \mathbf{D}_k^{\frac{1}{2}} \mathbf{U}_k^H \widehat{\mathbf{e}}_k \sim \mathbf{U}_k \mathbf{D}_k^{\frac{1}{2}} \mathbf{e}_k, \tag{3.2}$$

where $\widehat{\mathbf{e}}_k \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{0}_{N_T}, \mathbf{I}_{N_T}\right)$ and $\mathbf{e}_k \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{0}_r, \mathbf{I}_r\right)$. The singular values decomposition of the matrix $\mathbf{R}_k \in \mathbb{C}^{N_T \times N_T}$ is $\mathbf{R}_k = \mathbf{U}_k \mathbf{D}_k \mathbf{U}_k^H$, where $\mathbf{D}_k \in \mathbb{R}^{r \times r}$ is a diagonal matrix containing $r = \text{rank}(\mathbf{R}_k)$ strictly positive singular values of the matrix $\mathbf{R}_k$, and $\mathbf{U}_k \in \mathbb{C}^{N_T \times r}$ contains the corresponding singular vectors. rank(.) is the rank of a matrix, or the number of non zero singular values. The last part shows that the $\mathbf{h}_k$ and $\mathbf{U}_k \mathbf{D}_k^{\frac{1}{2}} \mathbf{e}_k$ distributions are the same.

In [4], a spatial channel correlation model is proposed. It assumes that the channels of the various antenna elements are Gaussian distributed, with zero mean and defined by a correlation matrix $\mathbf{R}$. The elements $r_{l,m}$ of the matrix $\mathbf{R}$ are defined as:

$$\begin{aligned} r_{l,m} &= \sum_{n=1}^{N_p} \mathbb{E}\left\{|g_n|^2\right\} \mathbb{E}\left\{e^{2\pi j d(l-1)\sin(\bar{\varphi})} e^{-2\pi j d(m-1)\sin(\bar{\varphi})}\right\} \\ &= \gamma \int e^{2\pi j d(l-m)\sin(\bar{\varphi})} f(\bar{\varphi}) d\bar{\varphi}. \end{aligned} \tag{3.3}$$

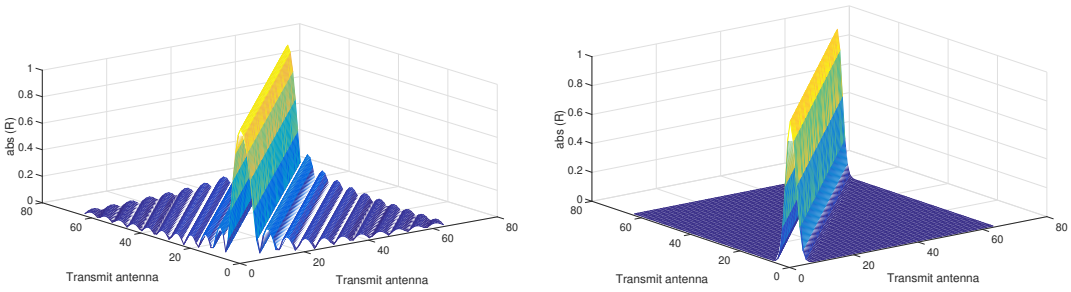
From here after, it is assumed that the multipath propagation is caused by a cluster of obstacles located around the users and that the base station is located above the obstacles, such that the line of sight propagation is not perturbed in an area around it. We denote $\bar{\varphi} = \varphi + \delta$, as the angle of each cluster, with φ being the deterministic nominal angle. δ is a random variable having a standard deviation from the nominal angles, σ_φ. When δ is uniformly distributed, $\delta \sim U(-\sqrt{3}\sigma_\varphi, \sqrt{3}\sigma_\varphi)$, the correlation model is called One Ring.

To reduce the complexity of the simulations, an approximation can be used for $r_{l,m}$ when σ_θ is small and δ follows a Gaussian distribution $\delta \sim \mathcal{N}(0, \sigma_\varphi^2)$, $r_{l,m}(\varphi) = \gamma e^{2\pi j d(l-m)\sin(\varphi)} e^{-\frac{\sigma_\varphi^2}{2}[2\pi d(l-m)\cos(\varphi)]^2}$.

The shape of $\mathbf{R}$ matrices are represented in Figure 3.1a and Figure 3.1b, for the One Ring and Gaussian correlation models, respectively.

3.3 Spatial compatibility

Two users are spatial compatible when their channels, $\mathbf{h}_k$, can be separated in space through linear precoding techniques. An equation that quantifies the amount



(a) One Ring correlation model (b) Gaussian correlation model

Figure 3.1: The shape of $\mathbf{R}$ matrix

of correlation between the channels of two users is described in [4, 5]. When $\mathbf{h}_i$ and $\mathbf{h}_k$ are the channels of two users, their inter-correlation can be defined as:

$$\cos\left(\angle\left(\mathbf{h}_i, \mathbf{h}_j\right)\right) \triangleq \mathbb{V}\left\{\frac{\mathrm{tr}\left[\mathbf{R}(\theta_i)\mathbf{R}(\theta_k)\right]}{N_T^2}\right\}, \tag{3.4}$$

where $\cos\left(\angle\left(\mathbf{h}_i, \mathbf{h}_k\right)\right) \in [0, 1]$.

4 Precoding algorithms

In this chapter, the main precoding schemes used in MU-MIMO system are presented. The total spectral efficiency and the computational complexity results studied in this chapter were published in [6], [7] and [8].

Linear precoding

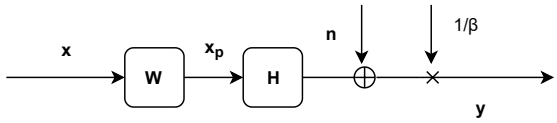
Linear precoding schemes for MU-MIMO work by projecting the useful signal of the user to an orthogonal sub-space of the other users [9], by inverting the channel matrix. The block scheme of a linear precoder is represented in Figure 4.1a.

Two of the linear precoding schemes, Zero Forcing (ZF) and Regularized Zero Forcing (RZF), are represented by the matrices $\mathbf{W} \in \mathcal{C}^{N_T \times K}$, $\mathbf{W}_{ZF} = \beta\mathbf{H}^{-1}$ and $\mathbf{W}_{RZF} = \beta\mathbf{H}^H\left(\mathbf{H}\mathbf{H}^H + \frac{N_T N_0}{E_x}\mathbf{I}\right)^{-1}$, where N_0 is the noise power spectral density, E_x is the energy of the transmitted signal and β is used to normalize the precoded signal in order to meet the total transmit power constraint.

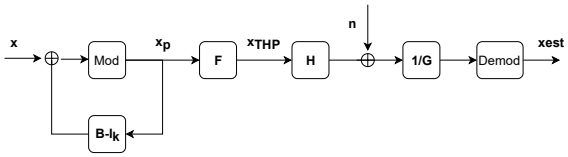
Block diagonalization of the channel matrix (BD) is a linear precoding scheme used for users with multiple receive antennas.

Non-linear precoding

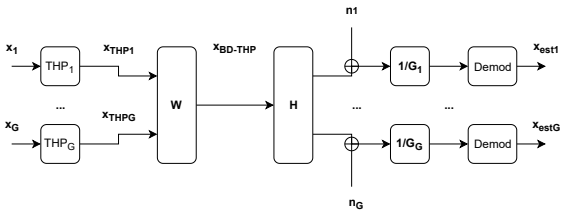
While linear precoding mechanisms have a relatively low computational complexity, their spectral efficiency is strongly affected in scenarios with a high density of users that exhibit strongly correlated channels. An alternative of those are the non-linear precoding schemes like Dirty Paper Coding (DPC) [10], Tomlinson Harashima Precoding (THP) and Vector Perturbation (VP).



(a) Linear precoder



(b) Non-linear THP precoder



(c) Hybrid BD-THP precoder

Figure 4.1: Block schemes

THP precoding scheme is illustrated in Figure 4.1b. The matrix $\mathbf{B} = [b_{kl}]$ is a unitary and lower triangular matrix, obtained through an LQ decomposition of the channel matrix $\mathbf{H}$ in an unitary matrix $\mathbf{F} \in C^{N_T \times K}$ and a lower triangular matrix $\mathbf{L} \in C^{K \times K}$, $\mathbf{H} = \mathbf{L}\mathbf{F}^H$. The equivalent channel matrix of the THP precoding scheme can be describer through the lower triangular matrix $B = \mathbf{G} \cdot \mathbf{H} \cdot \mathbf{F}$.

To prevent the total power growth during the precoding with $\mathbf{B}^{-1}$, a modulo function is used. The function ensures that the precoded symbols remain in the Voronoi region of the original signal constellation. In an M-Quadrature Amplitude Modulation (QAM) signal constellation, the modulo function extends periodically the constellation, by adding $2\sqrt{M}$ multiples to the imaginary and real parts of x_k.

Hybrid precoding

To reduce the impact of the channel estimation errors and the computational complexity of the non-linear precoding in multiple antenna systems, a hybrid, linear-non-linear precoding scheme was proposed in [11, 12, 13, 14]. In a first stage, a linear precoder is used to eliminate the interference between the groups of user with uncorrelated channels. In a second stage, a non-linear precoding scheme like THP is used to eliminate the interference between the users of the same group. Figure 4.1c illustrates the hybrid precoding scheme Block Diagonalization - Tomlinson Harashima Precoding (BD-THP).

4.1 Simulation results

Computational complexity evaluation

The computational complexity can be computed through the number of Floating Point Operations (FLOPs), required for multiplying and adding complex numbers. The computational complexity of the precoding schemes mentioned above is summarized in Table 4.1 [6, 12], where for the case of the hybrid precoder, $K_g = K/G$ represents the number of user in each group.

Table 4.1: Precoding schemes complexity

Precoding scheme	Number of FLOPs
RZF	$4K^3 + 2KN_T(4K-1) + K(4N_T-1)(K+1) + 2N_T(4K-1) + 8K^2 + 7K$
THP	$16K^3/3 + 2K(4KN_T - K + 2) + 2T(2K + 2K^2 + N_T(4K-1) - 4)$
BD-THP	$40GK_g^3/3 - 4GK_g^2 + 2GK_g - 2GK_gN_T + 16GK_g^2N_T + T(4GK_g^2 + 4GK_g + 8K_gN_T - 8G - 2N_T)$

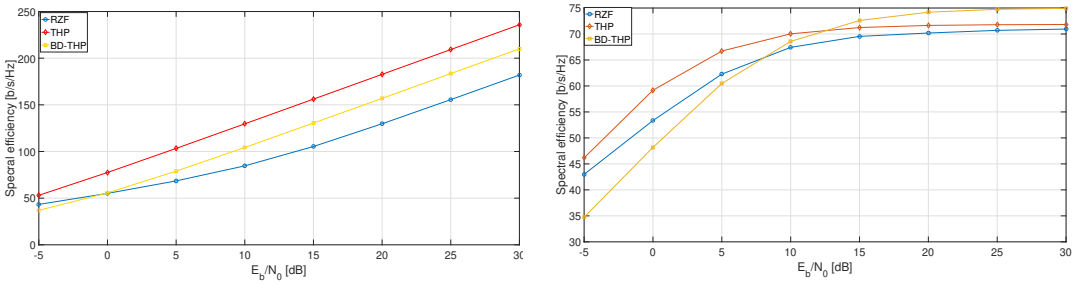
Spectral efficiency evaluation

The downlink of a single cell MU-MIMO system is considered. The base station is equipped with a linear array antenna having N_T elements uniformly distributed, that simultaneously transmit data to K single antenna users. The channel gains between each user and the N_T transmit antennas, follow a spatially correlated distribution.

The spectral efficiency is evaluated as a function of the energy per bit E_b, relative to the noise power spectral density N_0, $E_b/N_0 \triangleq E_x/\left[KN_0\log_2(M)\right]$, where M is the QAM modulation order. For the performance evaluation of the precoding schemes, Monte Carlo simulations were performed in a MU-MIMO system with $N_T = 16$ transmit antennas, $K = 16$ users and a $M = 16$-QAM modulation.

Figure 4.2a evaluates the performance of the precoding schemes when the channel is perfectly estimated and available at the base station, and Figure 4.2b, when the channel estimate is erroneous. Similar with [13], the error of the estimation is modeled as an Additive White Gaussian Noise (AWGN) of 25 dB variance, that corresponds to an user that moves at an approximate speed of 50 km/h, a carrier frequency of 2 GHz and a processing duration of 10 ms. Compared to Figure 4.2a, it can be noticed

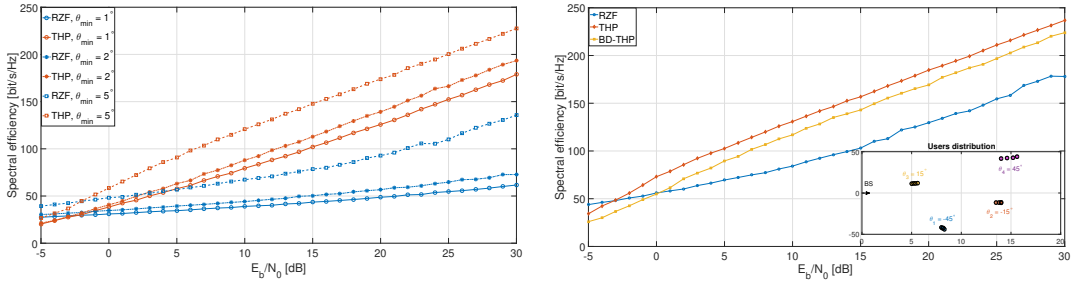
that BD-THP has lower performance than of the rest of the precoding schemes, as both BD and THP are impacted by the channel estimation error.



(a) Perfect channel estimation (b) Erroneous channel estimation

Figure 4.2: Spectral efficiency vs. E_b/N_0

Figure 4.3a shows the spectral efficiency of the RZF and THP precoding schemes, as a function of the E_b/N_0 and of the angle between users, when they are each separated by a fixed angle $\theta = \theta_{\min}$. The performance of both schemes decreases as the users get closer to each other, but the performance of the THP scheme is always better than that of the RZF precoding scheme, due to the successive interference cancellation procedure.



(a) $\theta_{\min} = 1°, 2°, 5°$ (b) $\theta_{\min,g} = 1°$ and $\theta_g = -45°, -15°, 15°, 45°$

Figure 4.3: Spectral efficiency vs. E_b/N_0 and $\theta_{\min}$

To evaluate the spectral efficiency of a hybrid BD-THP precoding scheme, in Figure 4.3b it is assumed that the users can be organized in four groups, each located at an angle $\theta_g = [-45°, -15°, 15°, 45°]$. As the minimum angle between any two groups is 30° and between any two users of the same group is 1°, a linear precoding scheme will be used between groups and a non-linear one in each group.

5 Users grouping

It was shown in the previous chapter that a hybrid precoding scheme can spatially multiplex users with correlated channel through non linear precoding techniques and

the users that can be spatially separated (those with uncorrelated channels), through linear precoding techniques. In this chapter, an algorithm for users grouping based on spatial correlation, is proposed. The results were published in [8] and [15].

5.1 K-Means algorithm

In order to optimize the performance of the users grouping method, the statistical K-means clustering algorithm [16] is proposed, using a distance metric based on the angles between users. The K-means algorithm groups together users such that each user to belong to the clusters with the closest center (the angle between the user and the center of the cluster is minimized).

This method guarantees that the users of the same cluster are close enough such that they cannot be multiplexed through a linear precoding scheme. In order for the users of two different clusters to be multiplexed through linear precoding, the angle between the centers of the clusters needs to be verified. A group of clusters whose users can be multiplexed through linear precoding should not include any cluster whose angle with the rest of the clusters in the group, is smaller than $\Delta\theta$, the minimum separation angle required by a linear precoder.

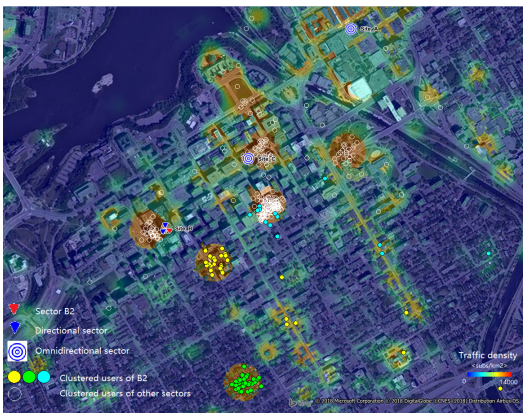
5.2 Simulation results

Evaluation of the clustering algorithm

For a fixed number of clusters and antennas, the proposed algorithm has a complexity of order $\mathcal{O}\left[K^{N+1} \times N\right]$[17], that is lower than of the two stage precoding scheme which uses the cordal distance metric, which is of order $\mathcal{O}\left[K^{N+1} \times N \times (N_T^2 + 2N_T^3)\right]$ [18]. The implementation of the K-means clustering algorithm through Lloyd's method can be optimized and exhibits a linear complexity with the number of users and clusters, of order $\mathcal{O}(KNI)$ [19]. It can be shown that the number of iterations required for convergence decreases a lot when the users already have a spatial distribution in clusters, which is the case of a realistic distribution.

In Figure 5.1, the users are distributed in two traffic scenarios, one which includes the morning and evening traffic (called Busy Hours), and the second, in the rest of day (called Working Hours). Due the movement of the users, the Busy Hours traffic is more distributed than at Working Hours, when is focused in hotspots, like the building offices. Using a minimum required angle between co-scheduled users of $\Delta\theta = 30°$, the users are grouped by the K-means algorithm in 4 clusters at Working Hours and in 3 clusters at Busy Hours. The two distributions were created in a radio planning software called Planet [20], using geolocated call traces [21, 22].

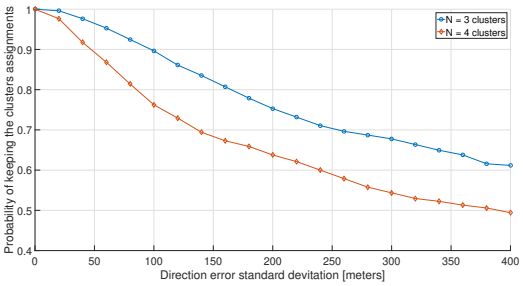
In practice, the direction of the users is not always perfectly estimated. For this reason, Figure 5.2 shows the impact of the estimation error of the directions of users on the algorithm's performance.



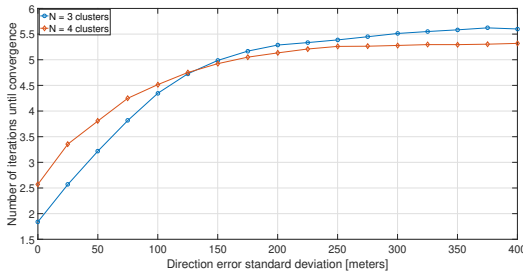
(a) Distribution and clustering of the users at Working Hours

(b) Distribution and clustering of the users at Busy Hours

Figure 5.1: Clustering results for B2 sector and $\Delta\theta = 30°$



(a) Probability of grouping the users in the same clusters



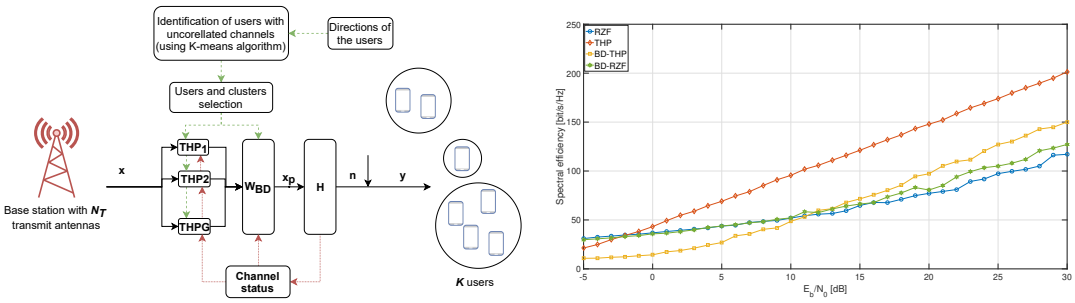
(b) Number of iterations until convergence

Figure 5.2: The performance on the clustering algorithm vs. the error standard deviation

Precoding scheme spectral efficiency evaluation

The scenario evaluated is represented in Figure 5.3a. It can be noticed that the clustering of the users requires knowing only the directions of the users and not their entire channel status. This reduces the required information for estimating and reporting the channel status information. Once the clustered are identified, a group of users is selected from every cluster in order to be served by a THP precoder. The groups of users selected and multiplexed through THP are then precoded through BD linear precoder. In the computation of the precoding weights, the full channel status information is only required for the selected users, and not for the entire set of users, making the algorithm compliant for Massive MIMO scenarios.

To evaluate the spectral efficiency (capacity) of a hybrid precoding scheme, in Figure 5.3b it is assumed that the users are distributed at Busy Hours. Compared to THP and RZF, the BD-THP hybrid precoding scheme has a similar performance



(a) Hybrid precoding scheme used together with the users clustering algorithm

(b) Spectral efficiency at Working Hours

Figure 5.3: Performance evaluation for the hybrid precoding scheme

as THP, the offset being caused by a lower performance of the BD linear precoding scheme. The new results of the hybrid precoding scheme being similar to those in the previous chapter, where the users were already organized in uncorrelated clusters, proves that the proposed algorithm organizes users for optimizing the performance of the precoding scheme.

6 Geolocation and spatial distribution of mobile network calls

In telecommunication networks, call traces containing information reported by the mobile equipments during calls, are collected automatically by the base stations for the purpose of investigating various events which may rise in the network and improving its performance. In this chapter, a technique for 3D geolocation of the LTE calls based on signal propagation predictions and a method for generating traffic spatial distributions, are presented. The results from this chapter were published in [21] and [22].

6.1 Implementation of the 3D geolocation algorithm

The main contribution of this chapter is the implementation of an enhanced algorithm to locate LTE calls in 3D, for indoor and outdoor environments, using information of the network configuration and LTE measurements. With **Loc** being the set of possible 3D locations, the probability of an LTE call to be located at location $\mathbf{Loc}_i = (X_i, Y_i, Z_i)$ when the set of reported measurements is $\mathbf{M}$, is computed as:

$$P(\mathbf{Loc}_i|\mathbf{M}) = P(\mathbf{Loc}_i|\mathbf{S}) \times P(\mathbf{Loc}_i|CellID) \times P(\mathbf{Loc}_i|TA), \tag{6.1}$$

where $\mathbf{S}$ is the set of Reference Signal Received Power (RSRP) measurements reported for $N \geq 1$ cells, TA is the value of the delay measured by the primary cell [Timing Advance (TA)] and $CellID$ is the unique cell identifier of the primary cell.

The location probability conditioned by the reported cell $P(\mathbf{Loc}_i|CellID)$, filters out the locations where the cell cannot be accessed (where its coverage is poor).

$P(\mathbf{Loc}_i|TA)$ filters out the locations too close or too far from the primary cell's location, using the value reported for the TA.

In order to locate a call, a filter based on the fingerprint of the measured received signals power is used, assuming the measured RSRPs are mutually independent and the error between predictions and measurements has a Gaussian distribution.

$$P(\mathbf{S}|\mathbf{Loc}_i) = \left\{ e^{-\frac{1}{2\sigma^2}\left(s_1^M - s_1^P\right)^2} \left[\prod_{j=2}^{N} e^{-\frac{1}{2\sigma^2}\left(\Delta s_j^M - \Delta s_j^P\right)^2} \right] \right\}^{\frac{1}{N}}, \tag{6.2}$$

where s_j^M, s_j^P are the measured RSRP values (present in the call traces) and those estimated using a propagation model, for the j cell, N is the number of reported RSRP values, and σ is a parameter that models the possible variations of the measured values compared to the estimated ones, due to fading or the measurement errors and $\Delta s_j^{M(P)} = s_1^{M(P)} - s_j^{M(P)}$.

In order to find the most probable location $\widehat{\mathbf{Loc}}$, two alternatives are possible:

- The location that has the maximum probability: $\widehat{\mathbf{Loc}} = \arg\max_{\mathbf{Loc}_i \in \mathbf{Loc}_s} P(\mathbf{Loc}_i|\mathbf{M})$;
- The location that minimizes the average location error: $\widehat{\mathbf{Loc}} = \frac{\sum_i \mathbf{Loc}_i \times P(\mathbf{Loc}_i|\mathbf{S})}{\sum_i P(\mathbf{Loc}_i|\mathbf{S})}$.

6.2 Measurements data base creation

In this section, a method for generating call traces with realistic measurements having valid 3D coordinates, is described. They are used to validate the performance of the 3D geolocation algorithm.

The steps required to generate the measurements starting from the predicted radio signals received power, are summarized as follows:

- Using the RSRP estimated for a number of cells in a random location, a noise modeling the fading effects is added;
- The fading is correlated between locations and between neighbor cells. The distances over which the fading is no longer correlated are of 7 meters indoors, and 30 meters outdoor.

6.3 Generation of the traffic distribution

Another contribution of this chapter is a method for generating traffic spatial distribution maps, using the probability values of the calls, computed for various locations by the geolocation algorithm, as well as the quantity of carried traffic by each call t_c:

$$T_{\mathbf{Loc}_i} = \frac{\sum_c t_c \times P(\mathbf{Loc}_i|\mathbf{S}_c)}{\sum_c P(\mathbf{Loc}_i|\mathbf{S}_c)}. \tag{6.3}$$

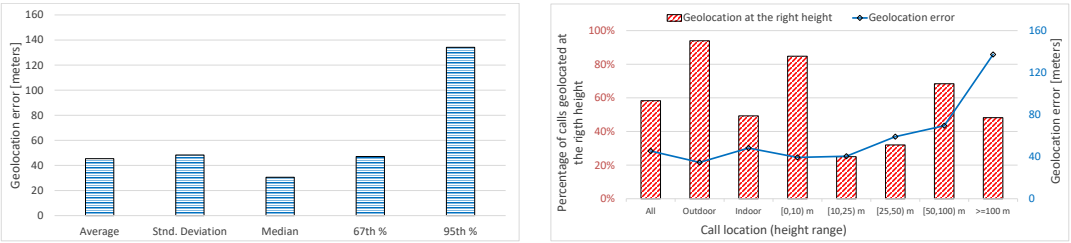
Compared to the existing methods, the one proposed does not require a precise location estimation.

6.4 Simulation results

The performance of the proposed 3D geolocation algorithm is evaluated using LTE calls generated around Tokyo train station. The Err precision is computed using the Euclidean distance between the real 3D coordinates (X_c, Y_c, Z_c) and those estimated, $(\widehat{X_c}, \widehat{Y_c}, \widehat{Z_c})$:

$$Err = \sqrt{\left(X_c - \widehat{X_c}\right)^2 + \left(Y_c - \widehat{Y_c}\right)^2 + \left(Z_c - \widehat{Z_c}\right)^2}. \tag{6.4}$$

In Figure 6.1a, it can be noticed that 95% of the calls are located with a precision of 140 meters and that the average location error is of 45 meters.



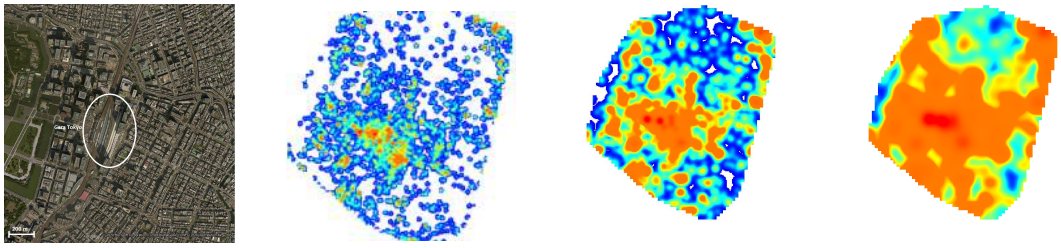
(a) In 3D (b) Per height

Figure 6.1: Geolocation algorithm performance

Figure 6.1b shows that the average location error increases with the height of the call. This is due to the performance of the geolocation algorithm based on fingerprinting, which deteriorates at higher heights because of the similarity of the estimated RSRP values.

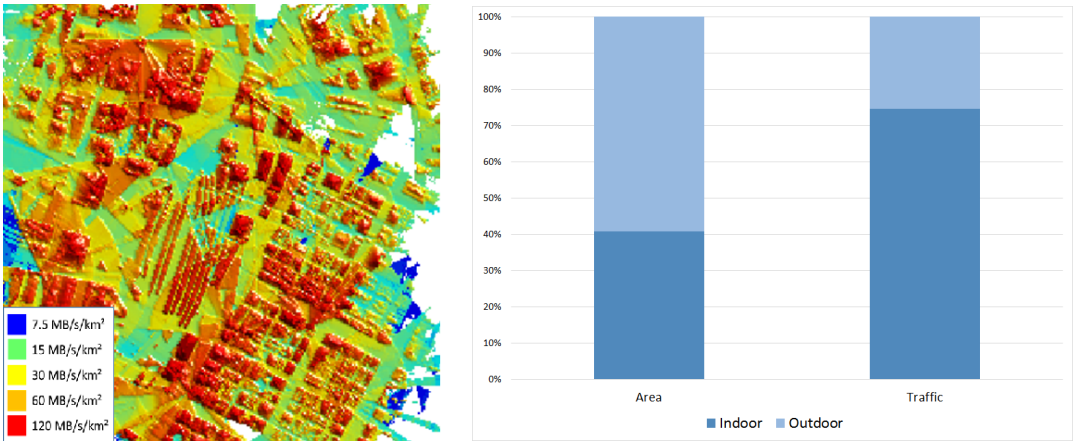
The performance of the method for generating the traffic distribution is evaluated using the geolocation algorithm on 208203 measurements of 134 cells. A visual analysis of the two methods indicates a better accuracy of the proposed method because:

- The building footprints are well distinguished in Figure 6.3a and match the footprints seen from the satellite (Figure 6.2a). In Figure 6.2, the contour of the buildings cannot be distinguished;
- The extremely busy area around Tokyo train station can be well distinguished, compared to the traffic distributions in Figure 6.2b, Figure 6.2c and Figure 6.2d, which indicate just a higher traffic density in that area.



(a) Satellite view of the simulated area (b) $\sigma = 5$ m (c) $\sigma = 15$ m (d) $\sigma = 30$ m

Figure 6.2: Traffic map generation using the standard method, a Gaussian spreading of traffic around the estimated locations



(a) Total traffic (b) Indoor and outdoor traffic

Figure 6.3: Traffic map generation using the proposed method

7 Simulation and tunning of the user scheduling algorithm

In this chapter, an algorithm for user selection and ordering is presented. It is based on the type of service, the channel quality and the accumulated bit rate. The users scheduling domain is not standardized and the techniques used are industrial secrets, as such, the algorithm proposed is conceived to be fully configurable, being able to be tuned for matching the performance of real base station implementation. The results presented in this chapter were published in [23] and patented in [24].

A presentation of the scheduling technique is made in [25]. In the category of those which account for the channel status and the Quality of Service (QoS), we find the techniques that work in Time Domain (TD) and in Frequency Domain (FD). The

performance of a two stage precoding algorithm (TD/FD) is also evaluated by the authors in [26, 27, 28]. More studies like [29], propose various metrics for the TD and FD algorithm.

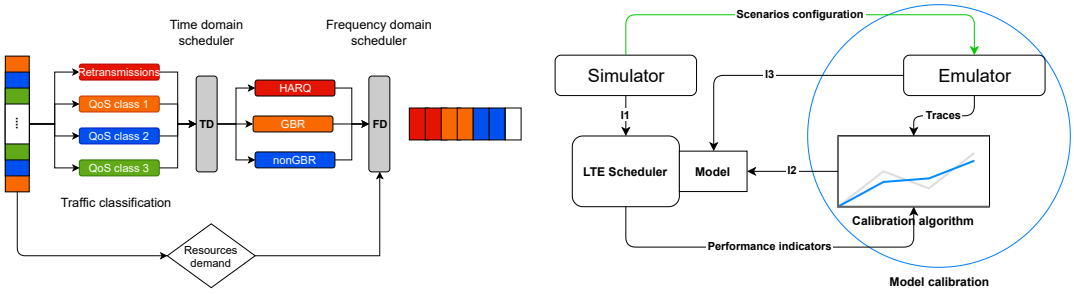
7.1 Scheduler model

A technique for modeling different types of traffic was conceived for the simulation of the scheduler behavior. This estimates the buffer length and the delay of the first packet, every time a transmission is possible, as a function of the average bit rate and the period of the packets. The traffic of the users with a Guaranteed Bit Rate (GBR) is periodic, whereas the traffic of the users with a non Guaranteed Bit Rate (non-GBR) is bursty.

Channel quality is described by the Signal-to-Interference plus Noise Ratio (SINR). It is assumed that the interference is proportional with the percentage of used Physical Resource Blocks (PRBs), or the load of each cell. The fast and frequency fading values are generated as a function of the channel type and the speed of the users.

At a specific time, the buffer contains either new packet, or packets that have not correctly been received by the users. The error probability of a packet is computed as a function of the SINR and the used Modulation and Coding Scheme (MCS).

The TD-FD scheduler was implemented to ensure the flexibility of the algorithm. It enables the algorithm to order users firstly in time, then to allocate the resources in frequency. The block scheme of the algorithm is depicted in Figure 7.1a.



(a) Scheduler algorithm (b) Calibration algorithm

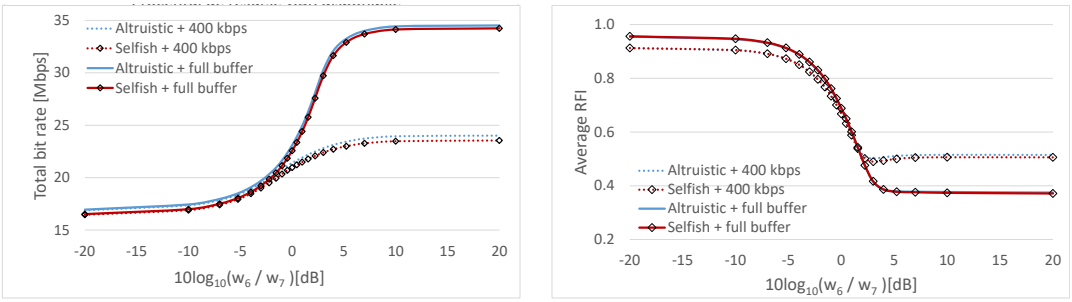
Figure 7.1: The block scheme of the system

In TD, the scheduling metric is defined to adapt to various QoS models, by ordering the users based on the buffer length, minimum guaranteed bit rate, accumulated bit rate, channel quality and service priority. The impact of each parameter is weighted by different configurable variables, w. In FD, the scheduling algorithm prioritizes firstly the resource allocation for retransmission, then for the GBR users and lastly, the available resources are given to non GBR users.

All the parameters that model the behavior of the scheduler, like the weights values, the number of resources available for GBR users, the altruistic or selfish strategy, enable the scheduler to be tuned to various implementations of real base stations.

7.2 Simulation results

The performance of various scheduling configurations is evaluated as a function of the total allocated bit rate and of the Rate Fairness Index (RFI), using the equation from [30].



(a) Total bit rate

(b) Rate Fairness Index

Figure 7.2: non GBR services performance vs. the radio of the channel quality to the accumulated bit rate, w_6/w_7

In Figure 7.2a and Figure 7.2b, when $10\log_{10}\left(\frac{w_6}{w_7}\right) = -20\text{dB}$, the scheduler serves firstly the users with the lowest accumulated bit rate, independent of their channel quality. When $10\log_{10}\left(\frac{w_6}{w_7}\right) = 20\text{dB}$, the total bit rate is maximized, but the RFI is very low.

Given the set of parameters of the scheduler mentioned above, in order to tune (calibrate) the scheduler algorithm, the architecture from Figure 7.1b is proposed. The simulation transmits through the I1 interference to the mobile equipment emulator and to the scheduler, the information regarding the number of users, their channel quality, the length of the scenario, the type of traffic, etc. Based on the preconfigured test scenarios, the emulator transmits through the I3 interface to the scheduler the values for the resources available for GBR, the maximum number of non GBR users fairly served, etc.

The calibration algorithm creates the list I2, by simulating the scheduler for various values of the w weights. Their optimal values minimize the difference between the

simulation and implementation and is found through a method based on the binary search principle.

8 Conclusions

8.1 Results

In the first chapter an introduction to the multiple antenna MIMO system used in cellular networks was made, and the purpose of the thesis was presented. In Chapter 2, a presentation of the 5G cellular systems was made, and the multiple antenna techniques were presented. In Chapter 3, the spatial and temporal channel correlation model was introduced, along with a way of estimating the spatial separation between users. In Chapter 4, the main precoding techniques were studies, which were then classified from a spectral efficiency (capacity) and computational complexity point of view. In Chapter 5, the users with a realistic distribution were grouped based on their spatial correlation. In Chapter 6, a method of generating the spatial distribution of users based on a 3D geolocation algorithm of LTE calls was presented. In Chapter 7, a method for user ordering and selection was presented, using a configurable radio resource scheduler.

8.2 Original contributions and publications

This section presents the main original contributions of this work, associated with the articles where they have been published.

Original contributions

(1) Summarizing the 5G NR specifications based on the initial drafts of the 3GPP RAN group, before the official publication in the standards.
(2), (3) Developing a performance testing system for the linear, non-linear and hybrid precoding schemes, a spatially correlated channel model, configurable using the angles between users and a method for estimating the spatial separation between users.
(4), (5) Developing a low complexity algorithm for grouping of users based on their spatial separation and adapting the hybrid precoding scheme based on the algorithm introduced.
(6), (7) Contributions to the development of a 3D geolocation algorithm of the LTE calls for testing the grouping algorithm and the precoding schemes, using a realistic traffic distribution.
(8) Contributions to the development of an adaptive users scheduler and of a method for evaluation of the measurements made by the mobile stations, in order to identify the behavior of as from an actual base station.

Publications

(1) R.-F. Trifan, A.-A. Enescu, and C. Paleologu, *Preview on MIMO Systems in 5G New Radio*, FABULOUS Proceedings, pp. 32-38, October 2017. [ISI Proceedings]

(2) R.-F. Trifan and C. Paleologu, *MU-MIMO Precoding Performance Conditioned by Inter-user Angular Separation*, ISETC Proceedings, pp. 1-4, November 2018. [ISI Proceedings]

(3) R.-F. Trifan și A.-A. Enescu, *Non-Linear Precoding Performance in Spatio-Temporally Correlated MU-MIMO Channels*, COMM Proceedings, pp. 181-186, May 2018. [ISI Proceedings]

(4) R.-F. Trifan, R. Lerbour, G. Donnard, and Y. L. Helloco, *K-Means MU-MIMO User Clustering for Optimized Precoding Performance*, VTC Proceedings, pp. 1-5, June 2019. [ISI Proceedings]

(5) R.-F. Trifan, A.-A. Enescu, and C. Paleologu, *Hybrid MU-MIMO Precoding Based on K-Means User Clustering*, Algorithms, vol. 12, pp. 1-18, July 2019. [ISI-Q3, FI 1,510]

(6) R.-F. Trifan, *Hybrid MU-MIMO Precoding Based on K-Means User Clustering*, SAD - ETTI, July 2019.

(7) R.-F. Trifan, R. Lerbour, and Y. L. Helloco, *Enhanced 3D Geolocation Algorithm for LTE Call Traces*, VTC Proceedings, pp. 1-5, September 2016. [ISI Proceedings]

(8) R. Lerbour, Y. L. Helloco, and R.-F. Trifan, *Hotspot Identification through Call Trace Analysis*, VTC Proceedings, pp. 1-5, September 2016. [ISI Proceedings]

(9) R. Lerbour and R.-F. Trifan, *Devices and Method for Simulating a Mobile Telecommunications Network*, European Patent Office, October 2016. [Patent]

(10) R.-F. Trifan, R. Lerbour, and Y. L. Helloco, *Mirroring LTE Scheduler Performance with an Adaptive Simulation Model*, VTC Proceedings, pp. 1-5, May 2015. [ISI Proceedings]

8.3 Perspectives of future development

A first perspective can be the integration of beamforming in the MIMO channel modeling, in order to accurately test the grouping algorithm and the precoding performance. Another perspective of development can be the use of the artificial intelligence algorithms to create an universal precoding scheme that can optimize some performance indicators, for various channel conditions and traffic distributions. The proposed algorithms could finally be embedded on a hardware architecture, in order to validate their real-time functioning.

Bibliography

[1] R.-F. Trifan, A.-A. Enescu, and C. Paleologu, "Preview on MIMO Systems in 5G New Radio," *FABULOUS Proceedings*, pp. 32–38, July 2017.

[2] T. Chien and E. Björnson, "Massive MIMO Communications," *5G Mobile Communications*, pp. 77–116, 2017.

[3] E. Bjornson, J. Hoydis, and L. Sanguinetti, "Massive MIMO has Unlimited Capacity," *IEEE Transactions on Wireless Communications*, vol. 17, pp. 574–590, January 2018.

[4] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency," *Foundations and Trends in Signal Processing*, vol. 11, pp. 154–655, November 2017.

[5] M. Herdin, N. Czink, H. Ozcelik, and E. Bonek, "Correlation Matrix Distance, a Meaningful Measure for Evaluation of Non-Stationary MIMO Channels," *VTC Proceedings*, pp. 136–140, December 2005.

[6] R.-F. Trifan and C. Paleologu, "Non-Linear Precoding Performance in Spatio-Temporally Correlated MU-MIMO Channels," *COMM Proceedings*, pp. 181–186, May 2018.

[7] R.-F. Trifan and A.-A. Enescu, "MU-MIMO Precoding Performance Conditioned by Inter-user Angular Separation," *ISETC Proceedings*, pp. 1–4, November 2018.

[8] R.-F. Trifan, A.-A. Enescu, and C. Paleologu, "Hybrid MU-MIMO Precoding Based on K-Means User Clustering," *Algorithms*, vol. 12, pp. 1–18, July 2019.

[9] E. Castaneda, A. Silva, A. Gameiro, and M. Kountouris, "An Overview on Resource Allocation Techniques for MU-MIMO Systems," *IEEE Communications Surveys & Tutorials*, vol. 19, pp. 239–284, Martie 2016.

[10] M. Costa, "Writing on Dirty Paper," *IEEE Transactions on Information Theory*, vol. 29, pp. 439–441, May 1983.

[11] F. Hasegawa, H. Nishimoto, N. Song, M. Enescu, A. Taira, A. Okazaki, and A. Okamura, "Non-Linear Precoding for 5G NR," *CSCN Proceedings*, pp. 1–7, October 2018.

[12] S. Zarei, W. Gerstacker, and R. Schober, "Low Complexity Hybrid Linear/Tomlinson-Harashima Precoding for Downlink Large-Scale MU-MIMO Systems," *Globecom Proceedings*, pp. 1–7, December 2016.

[13] M. Electric, "MU-MIMO Performance Evaluation of Nonlinear Precoding Schemes," *Raportul întâlnirii 3GPP TSG RAN WG1 88b*, pp. 1–5, April 2017.

[14] Huawei and HiSilicon, "Non-linear Precoding for Downlink MU-MIMO," *Raportul întâlnirii 3GPP TSG RAN WG1 88*, pp. 1–5, February 2017.

[15] R.-F. Trifan, R. Lerbour, G. Donnard, and Y. L. Helloco, "K-Means MU-MIMO User Clustering for Optimized Precoding Performance," *VTC Proceedings*, pp. 1–5, June 2019.

[16] S. Lloyd, "Least squares quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28, pp. 129–137, Martie 1982.

[17] M.Inaba, N. Katoh, and H. IMay, "Applications of Weighted Voronoi Diagrams and Randomization to Variance-based K-clustering," *10th ACM Symposium on Computational Geometry Proceedings*, p. 332–339, June 1994.

[18] Y. Xu, G. Yue, and S. Mao, "User Grouping for Massive MIMO in FDD Systems: New Design Methods and Analysis," *IEEE Access*, vol. 2, pp. 947–959, August 2014.

[19] G. Hamerly, "Making K-means Even Faster," *SIAM International Conference on Data Mining Proceedings*, pp. 130–140, May 2010.

[20] InfoVista, "Planet, software de planing radio." `https://www.infovista.com/planet/rf-planning-software`. [accesat pe 08 May 2020].

[21] R. Lerbour, Y. L. Helloco, and R.-F. Trifan, "Hotspot Identification through Call Trace Analysis," *VTC Proceedings*, pp. 1–5, September 2016.

[22] R.-F. Trifan, R. Lerbour, and Y. L. Helloco, "Enhanced 3D Geolocation Algorithm for LTE Call Traces," *VTC Proceedings*, pp. 1–5, September 2016.

[23] R.-F. Trifan, R. Lerbour, and Y. L. Helloco, "Mirroring LTE Scheduler Performance with an Adaptive Simulation Model," *VTC Proceedings*, pp. 1–5, May 2015.

[24] R.-F. Trifan and R. Lerbour, "Devices and Method for Simulating a Mobile Telecommunications Network - EP3526932A1," *European Patent Office*, October 2016.

[25] T. K. Ramesh, "A Survey on Scheduling Algorithms for Downlink in LTE Cellular Network," *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 219–223, February 2019.

[26] W. Fu, Qingliang, W. Tian, C. Wang, and L. Ma, "A QoS Aware Scheduling Algorithm Based on Service Type for LTE Downlink," *VTC Proceedings*, pp. 2468–2473, Martie 2013.

[27] B. Bojovic and N. Baldo, "A New Channel and QoS Aware Scheduler to Enhance the Capacity of Voice Over LTE Systems," *SSD Proceedings*, pp. 1–6, February 2014.

[28] Y. Zaki, T. Weerawardane, C. Gorg, and A. Timm-Giel, "Multi-QoS aware Fair Scheduling for LTE," *VTC Proceedings*, pp. 1–5, May 2011.

[29] M. M. Nasralla and I. U. Rehman, "QCI and QoS Aware Downlink Packet Scheduling Algorithms for Multi-Traffic Classes over 4G and beyond Wireless Networks," *3ICT Proceedings*, pp. 1–7, November 2018.

[30] R. Jain, D. Chiu, and W. Have, "A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer Systems," *DEC Research Report*, pp. 1–37, September 1984.