Comparison Study of Scheduling Strategies for Multi-user MIMO OFDM Systems with Limited feedback

Krishna .P, Kishan Rao.K, and Anil kumar.T

***Abstract*—** Scheduling schemes provide efficient support of multimedia services in multi-user MIMO OFDM wireless networks. In cellular systems the base station transmits a signal intended for single user in a particular resource allocation. The efficient utilization of available bandwidth in a system essential in modem wireless systems such as video streaming and voice over internet protocol(VoIP), which demands for higher data rates. Since the user feedback the channel state information in the network, base station to schedule the more than one user data in a single resource allocation. For properly utilizing the feedback information, scheduling algorithms are designed which select pairs of users which would maximize system capacity. In this paper, We describe and compare the scheduling algorithms to improve the system capacity and fairness among the users proposed for the 3GPP Long term evolution (LTE) systems**.**

***Keywords*—**MU-MIMO, Bandwidth, Limited Feedback,3GPP-LTE.

I. Introduction

As the demand for high data rate application like video and audio streaming, VoIP, video conferencing are increasing, future wireless systems should be able to provide high speed broadband services for mobile users with sufficient quality of services (QoS) support. As the bandwidth and power are scarce or limited resources, techniques which lead to efficient utilization of these resources are quite necessary in the next generation wireless systems. At the same time the wireless channel creates a challenging environment because of variety of channel impairments. Thus, future wireless systems are to be designed taking all these factors into consideration.

MIMO systems which employ multiple antennas at transmitter and receiver is a very useful technique in wireless environments to combat the effect of fading and to use the radio channel efficiently by transmitting multiple streams to a user in the same resource allocation, thus achieving diversity gain and multiplexing gain, which is the first in achieving high spectral efficiencies [1]-[3].

Krishna.Patteti is with the SVS Institute of Technology, Hanamkonda, Warangal, 506015, INIDA (corresponding author's phone: 09963309062; e-mail:kpatteti@gmail.com).

Kishan Rao.K was with Vaagdevi College of Engineering, Warangal, AP 506005, INDIA. He is now with the Department of ECE, (e-mail: prof_kkr@rediffmail.com).

Anil Kumar Tipparthi is with the ECE Department, SR Engineering College, Warangal,AP, 506371, INDIA (e-mail:tvakumar2000@yahoo.co.in).

Orthogonal frequency division multiplexing (OFDM) has long been regarded as an efficient approach to mitigate the effect of inter-symbol interference (ISI) in frequency-selective channels by dividing the entire channel into many narrow parallel sub-channels [4].Therefore the combination of these two technologies, which is called MIMO-OFDM, is an efficient way for providing high data rate reliable communications [5].

For a large number of users to be served in one cell, high capacity gains can be achieved by transmitting independent data streams to different users sharing the same time-frequency resources. This technique is referred to multi-user multiple-input multiple-output (MU-MIMO) [2].It is one of the techniques which can be used in cellular systems to increase spectral efficiency. In MU-MIMO operation two or more user environment's (UE) or mobiles share the same time frequency resources. Several parallel data streams are transmitted simultaneously, one for each UE. It is assumed that the UE feedback a quantized version of the observed channel, so that the base station (BS) can schedule in MU-MIMO mode terminal with good channel separation.

II. Overview of LTE FRAME STRUCTURE and CQI Feedback

In LTE, ten 1 ms sub frames compose a 10 ms frame. Each sub frame divides into two slots. The smallest modulation structure in LTE is the Resource Element. A Resource Element is one 15 kHz sub carrier by one symbol. Resource Elements aggregate into Resource Blocks. A Resource Block has dimensions of sub carriers by symbols. Twelve consecutive sub carriers in the frequency domain and six or seven symbols in the time domain form each Resource Block.

The number of symbols depends on the Cyclic Prefix (CP) in use. When a normal CP is used, the Resource Block contains seven symbols. When an extended CP is used, the Resource Block contains six symbols. A delay spread that exceeds the normal CP length indicates the use of extended CP.

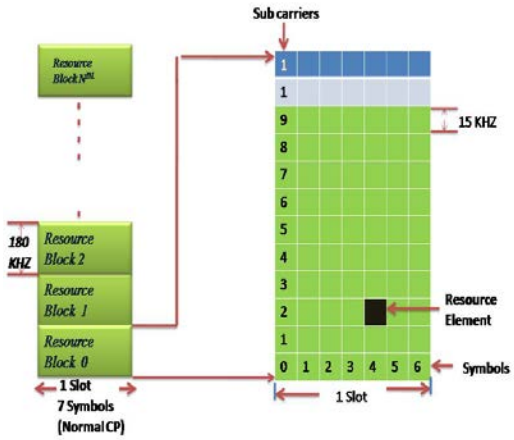


Fig:1 Relationship between a slot, symbols and Resource Blocks. N^{DL} is the symbol used to indicate the maximum number of downlink Resource Blocks for a given bandwidth.

In the time domain the LTE transmissions are organized into frames of 10 msec length. Each frame is composed of 10 sub frames of 1 msec duration. Each sub frame is made up of two equal sized slots of 0.5 msec each. Each slot is composed of 7 or 6 OFDM symbols depending upon whether a short or long Cyclic Prefix (CP) has been used. In the case of a short CP the CP has duration of 5.21µsec for the first symbol and 4.69 µsec for the remaining 6 symbols such that the total slot duration is 0.5 msec. For the long CP the CP has duration of 16.67 µsec for all 6 symbols. The useful symbol duration is fixed at 66.67 µsec. In the frequency domain the subcarriers are spaced at 15 kHz giving a useful symbol duration of 1/15000=66.67µsec after the IFFT operation. A group of 12 subcarriers (12*15=180 kHz) over one time slot (0.5 msec) is referred to as a Resource Block (RB). With 1200 subcarriers available over a 20 MHz bandwidth the LTE transmitter has 100 RBs available over one time slot. The smallest unit of resource is termed as the Resource Element which is composed of one subcarrier and one OFDM symbol. Thus an RB has 12×7=84 and 12×6=72 Resource Elements for short and long CP respectively. *Channel Quality Indicator (CQI) Feedback:* The CQI is a 4-bit value that indicates an estimate of the modulation and coding scheme (MCS) that the UE can receive reliably from the BS. It is typically based on the measured received signal quality, which can be estimated, for example, using the pilots sent by the BS on the downlink.

III. SYSTEM MODEL

We consider the downlink of a wireless cellular system based on OFDM with Nc sub carriers. The BS has M transmitting antennas and N receiving antennas. If U user are waiting to be scheduling at the BS, the scheduler will determine K(<U) users according to scheduling algorithm and feedback sent by the MSs. At the same time scheduler will select a precoding matrix W from the codebook to precode for the scheduled K user before transmission. The received signal at the k^{th} MS is given by:

$$y_k = H_k W x_k + n_k \quad k=1,2,\ldots\ldots U \quad (1)$$

Where

$$x_k = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_k \end{bmatrix}_{k\times 1}$$

$$H_k = \begin{bmatrix} h^k{}_{11} & h^k{}_{1,2} & . & . & h^k{}_{1M} \\ h^k{}_{2,1} & h^k{}_{22} & . & . & h^k{}_{2M} \\ . & . & . & . & . \\ . & . & . & . & . \\ h^k{}_{N,1} & . & . & . & h^k{}_{NM} \end{bmatrix}_{N\times M}$$

and

$$W = \begin{bmatrix} v_1 \\ v_2 \\ . \\ . \\ v_k \end{bmatrix}_{k\times 1}$$

.xk is the complex symbol transmitted for the kth user, Hk is the NXM wireless matrix from the kth MS to BS. W is the precoding matrix from the codebook to pre code the scheduled K users before transmission and nk is the noise vector at the kth MS.

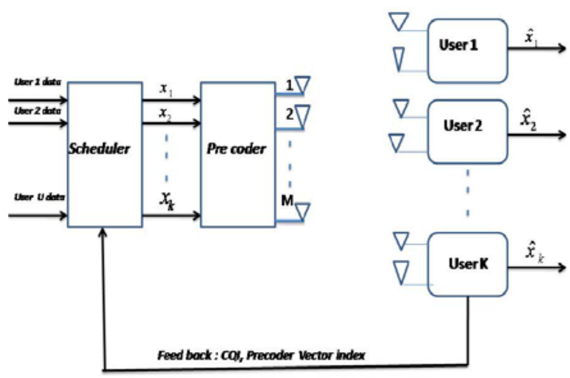


Fig. 2 MU-MIMO System

A. *MMSE Receiver:*

For traditional MIMO detection a linear receiver is used to detect the transmit data. ZF (zero forcing), MMSE and MRC detection criterions are commonly employed. In order to obtain good performance, We consider a linear MMSE receiver will try to reduce the MSE between the desired and estimated symbols. The linear 1X N MMSE receiver b1 to detect 1^{st} user data can be expressed as:

$$b_1 = \min_b E\left\{\left|x_1 - \hat{x}_1\right|^2\right\} = \min_b E\left\{\left|x_1 - y_1\right|^2\right\} = p^* R^{-1}$$
(2)

Where R is $N \times N$ auto correlation matrix of received vector y_1, p is $N \times 1$ cross-correlation vector between the desired symbol x_1 and received vector y_1. Assuming the total power constraint P is divided equally among K users, expression for R and p is calculated as follows:

$$R = E\{y_1 y_1^*\}$$
$$= \tilde{H}_1 E\{xx^*\}\tilde{H}^*_1 + E\{n_1 n_1^*\}$$
$$= \left(\frac{P}{K}\right)\tilde{H}_1 H_1^* + N_0 I_N \quad (3)$$

$$p = E\{y_1 x_1^*\}$$
$$= H_1 v_1 E\{x_1 x_1^*\}$$
$$= \left(\frac{P}{K}\right) H_1 v_1$$
$$= \left(\frac{P}{K}\right)\tilde{h}_1 \quad (4)$$

Where $\tilde{h}_1 = H_1 v_1 = \left|\tilde{H}_1\right| = \left|H_1 W_1\right|$ and * indicates conjugate transpose operation. The final expression for MMSE filter is written as:

$$b_1 = \tilde{h}_1^* \left[\tilde{H}_1 \tilde{H}_1^* + \frac{KN_0}{P} I_N\right]^{-1} \quad (5)$$

B. Feedback Mechanism:

Feedback information is sent every frame. Feedback at a subscriber level is not feasible since it would require a high capacity feedback link. We use sub band level feedback. Where each user returns 1) PVI and 2) CQI information for every subband.This PVI selection is made such that it represents the best choice of precoding vector from the precoder used for the given sub band. The CQI estimate that is feedback is averaged over all sub bands in sub band.

C. CQI Calculations

Let us calculate the CQI of the 1st user. The estimated symbol at the k^{th} MS can be written as:

$$\hat{x} = b_1 y_1$$
$$= b_1 H_1 v_1 x_1 + \sum_{i=1}^{K} b_1 H_1 v_i x_i + b_1 n_1 \quad (6)$$

Where b_1 is MMSE receiver vector. Assuming the total power P is divided equally among all K users, expression for signal(S), interference (I) and noise power (N) of 1st user is calculated as follows:

$$S = E\{(b_1 H_1 v_1 x_1)(b_1 H_1 v_1 x_1)^*\}$$
$$= \left|b_1 H_1 v_1\right|^2 E\{x_1 x_1^*\}$$
$$= \left(\frac{P}{K}\right)\left|b_1 H_1 v_1\right|^2 \quad (7)$$

$$I = \sum_{i=1}^{K} E\{(b_1 H_1 v_i x_i)(b_1 H_1 v_i x_i)^*\}$$
$$= E\{x_1 x_1^*\} \sum_{i=1}^{K} \left|b_1 H_1 v_i x_i\right|^2$$
$$= \left(\frac{P}{K}\right) \sum_{i=1}^{K} \left|b_1 H_1 v_i\right|^2 \quad (8)$$

$$N = E\{(b_1 n_1)(b_1 n_1)^*\}$$
$$= b_1 E\{b_1 n_1^*\} b_1^*$$
$$= \left\|b_1\right\|^2 N_0 \quad (9)$$

$$CQI_1 = \frac{S}{I + N} \quad (10)$$

Here $\|.\|$ indicates norm of the vector. Since noise is assumed to be Gaussian distributed.

Capacity of 1st user C_1 can be calculated from shannon's capacity theorem as:

$$C_1 = \log_2(1 + CQI_1) \quad (11)$$

After calculating all individual capacities, the sum capacity can be calculated by adding individual capacities.

$$C_{sum} = \sum_{i=1}^{K} C_i \quad (12)$$

D. Scheduling Algorithms

In wireless networks classical access method like Round Rabin(RR) and Random access (RA) are not well suited and provide poor throughput. Consequently much interest has recently been given to the design of scheduling schemes that maximize the performance of multiuser OFDM systems. In the following, we focus on the three major scheduling techniques that emerged: Round Robin algorithm (RR), Maximum Signal-to-Noise Ratio (MaxSNR), Proportional Fair (PF) scheduling and BABS algorithm.

1) Round Robin Algorithm:

Round Robin (RR) is one of the simplest scheduling algorithms, which assigns time slices to each MS in equally portions and in order, handling all MSs as having the same priority. RR scheduling is both simple and easy to implement, and starvation-free.

The main reason for this algorithm does not take into account the changing reception conditions at the different receivers .RR algorithm does not take into account multiuser diversity and thus it will schedule transmissions to/from subscribers half of the time when their reception conditions are worse than average [7].

2) Maximum Signal-to-Noise Ratio Scheduling: Many schemes are derived from the Maximum Signal –to- Noise Ratio (MaxSNR) technique which allocates resource at a given time to active mobiles with greatest SNR. Denoting $m_{k,n}$ the maximum number of bits that can be transmitted on a time slot of resource unit n if it is allocated to the mobile k, MaxSNR scheduling consists in allocating the Resource Unit n to the mobile j which has the greatest $m_{k,n}$ such as

$$j = \arg\max_{k}\left(m_{k,n}\right) \qquad k = 1,2\ldots\ldots..U \qquad (13)$$

Where U is the total number of mobiles. Profiting of multiuser and bandwidth granularity, MaxSNR continuously allocates the radio resources to the mobiles with the best spectral efficiency. However MaxSNR assumes that the user with the most favorable transmission conditions has information to transmit at the considered time instant. It does not take into account the variability of the traffic and the queuing aspects. Additionally, a negative side effect of this strategy is that the closest mobiles to the access point have disproportionate priorities over mobiles more distant since their path loss attenuation is much smaller. This results in a severe lack of fairness.

3) Proportional Fair (PF) scheduling:

The PF scheduler is designed to take advantage of multiuser diversity while maintaining comparable long-term throughput for all users. The basic principle is to allocate resources to a mobile j when its channel conditions are most favorable with respect its time average such as

$$j = \arg\max_{k}\left(\frac{m_{k,n}}{M_{k,n}}\right) \qquad k = 1,2,\ldots\ldots U \qquad (14)$$

Where $M_{k,n}$ is the time average of the $m_{k,n}$ value. The average throughput $M_{k,n}(t)$ for all users is then updated according to

$$M_{k,n}(t+1) = \left(1 - \frac{1}{t_c}\right) M_{k,n}(t) + \frac{1}{t_c} m_{k,n} \qquad k = k^*$$

$$= \left(1 - \frac{1}{t_c}\right) M_{k,n}(t) \qquad k \neq k^* \qquad (15)$$

$$\text{PFMetric} = \frac{m_{k,n}}{M_{k,n}} \qquad (16)$$

Thus consistently underserved users receive scheduling priority, which promotes fairness. The parameter t_c controls the latency of the system. If t_c is large, the latency increases, with benefit of higher sum throughput. If t_c is small, the latency decreases, since the average throughput values change more quickly, at expense of some throughput.

At a short time scale, path loss variations are negligible and channel states in formations are mainly due to multipath fading, statistically similar for all mobiles. Thus, PF provides an equal sharing of the total available bandwidth among the mobiles as RR. Actually, PF combines the advantages of the classical schemes and the opportunistic schemes. It currently appears as the best bandwidth management scheme.

In PF-based schemes, fairness consists in guaranteeing an equal share of the total available bandwidth to each mobile, whatever its position or channel conditions. However, since the farther mobiles have a lower spectral efficiency than the closer ones due to path loss, all mobiles do not all benefit of an equal average throughput despite they all obtain an equal share of bandwidth.

In particular PF scheduling does not take into account the delay constraints and is not well adapted to multimedia services which introduce heterogeneous users, new traffic patterns with highly variable bit rates and stringent QoS requirements in terms of delay and packet loss.

4) BABS Algorithm: Band Assignment based on SNR (BABS) algorithm.

The implemented BABS algorithm works as follow: At first each user get mk subcarriers according to their rate requirements, if the sum of the whole sub-carriers allocated does not fill with the available bandwidth, we have to

1. Remove sub-carriers from users who are demanding fewer sub-carriers.

2. Add sub-carriers to users according to the SNR.

In order to make relevant comparisons with the other algorithms, fix power at the BS is used .So the BS simply distributes the transmission power equally among the sub-carriers without taking advantage on power control.

In [8], this algorithm is shown to converge to the distribution of sub- carriers among users and the function f(r) is explained more accurately.

BABS Algorithm flow chart:

$m_k \leftarrow \left\lceil \frac{R_{min}^k}{R_{max}} \right\rceil$ k=0…..K-1.

While

$\sum_{k=0}^{K-1} m_k > N$ do

$k^* \leftarrow argmin_{0 \le k \le K-1} m_k$

End while

While
$\sum_{k=0}^{K-1} m_k < N$ do
$G_k \leftarrow \frac{m_k+1}{H_k} + \left\lceil \frac{R_{min}^k}{m_k+1} \right\rceil - \frac{m_k}{H_k} f\left(\frac{R_{min}^k}{m_k}\right)$
1$l \leftarrow argmin_{0 \le k \le K-1} G_k$
$m_l \leftarrow m_l + 1$
end while.

IV. Results Anyasis

All the users simulated have the same channel response average (where users have equal channel responses average and have been simulated according to filtered Gaussian noise method). In simulations there are 16 sub-carriers for the OFDM system and the carrier frequency is 2.4 GHz. In order to show how work the different scheduling algorithms we only simulate four different users, A (red), B (blue), C (green) and D (yellow), with their different channel responses. We assume that user A is transmitting video (64 Kb per time slot), user B is transmitting voice (16 Kb per time slot), users C and D are transmitting data, 30 and 40 Kb per time slot respectively. Each sub-carrier is allocated to each user in two different scenarios according to four different algorithms, RR, Max Rate, PFS and RCG. Table I: shows simulation parameters.

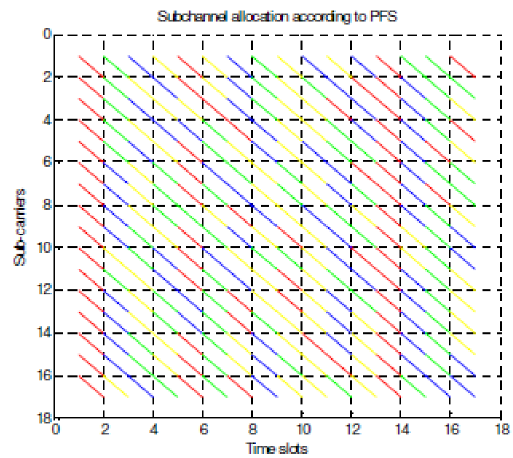
Table I
Simulation Parameters

Parameters	Value
Time Slots	18
Sub-Carriers	18
Carrier Frequency(Gzh)	2.4
Average Throughput	20
Users	6
Rate kp per timeslots	64
Mobile Speed	5

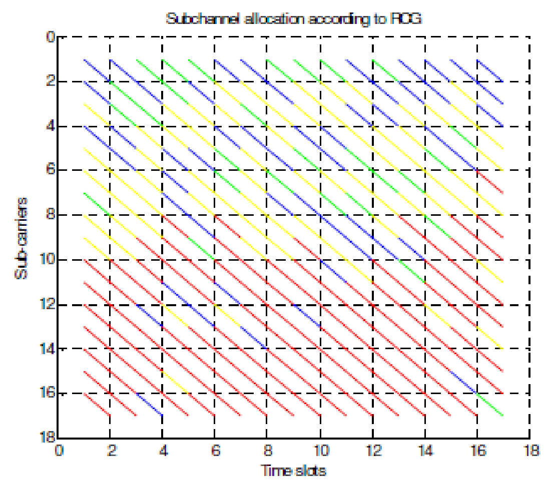
RR algorithm allocates all sub-carries to one user at each time slot, so this policy is totally fair. In this scenario we cannot notice the differences between max.SNR (see.Fig3) and PF scheduling (Fig 3) policies on allocating resources because all users have the same channel response average. But we can notice that RCG (BABS) allocates more sub-carrier to users. Than other scheduling algorithms, reason is because only this algorithm takes into account rate requirements by users.

A.Scenario A: Equal Users

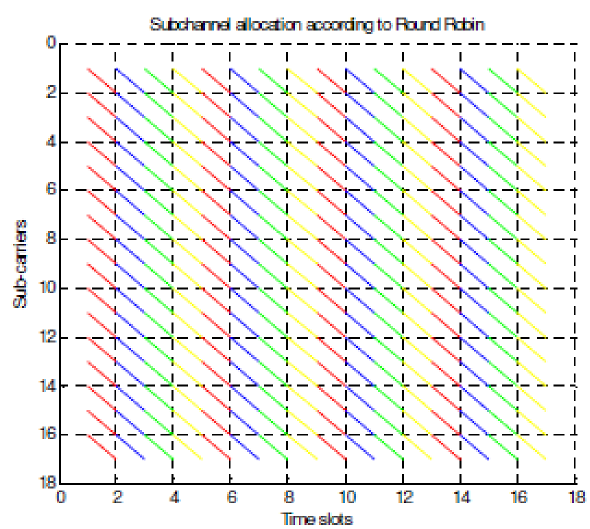
As we can see in Fig3, all the users simulated have the same channel response average. In Fig3, the scheduler scheme is showed according to four different scheduling policies. As we can see in Fig3 (a), RR algorithm allocates all sub-carriers to one user at each time slot, so this policy is totally fair. In this scenario we cannot notice the differences between Max Rate (see Fig3.(b) and PFS (see Fig3(c) policies on allocating resources because all users have the same channel response average. But we can notice that RCG allocates more sub-carriers to user A than other scheduling algorithms, the reason is because only this algorithm takes into account rate requirements by users and user A is transmitting video and this services needs more amount of data than other services, so user A needs more sub-carriers than other users.



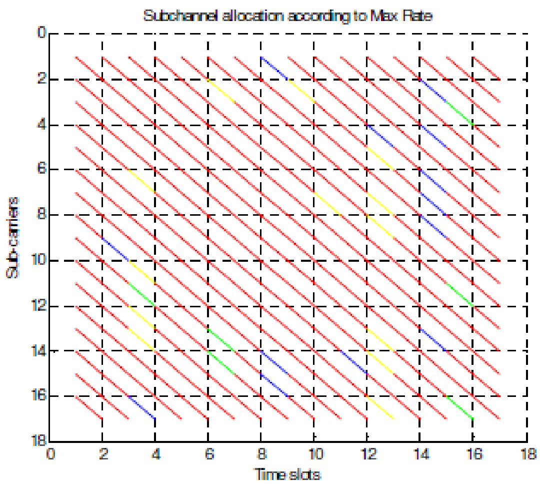
(a)



(b)



(c)

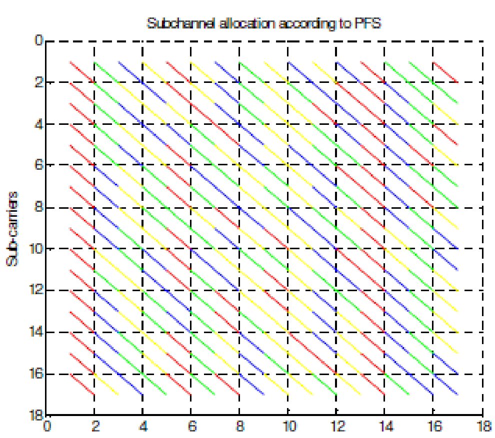


(d)

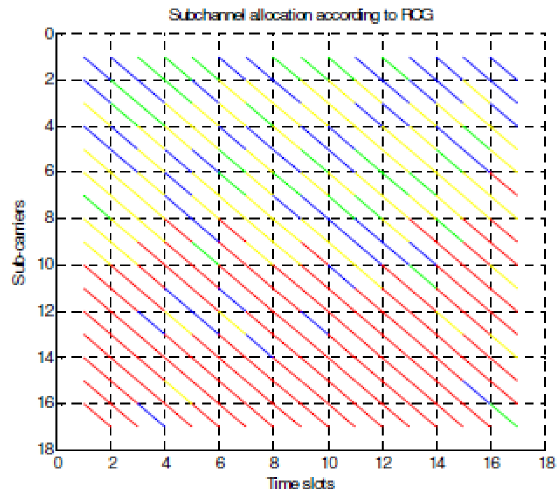
Fig. 3 Scheme where users have equal channel responses average and have been simulated according to filtered Gaussian noise method, (a) RR algorithm, (b) Max Rate algorithm, (c) PFS algorithm, with tc =18, (d) RCG algorithm.

B.Scenario B: Un- Equal Users

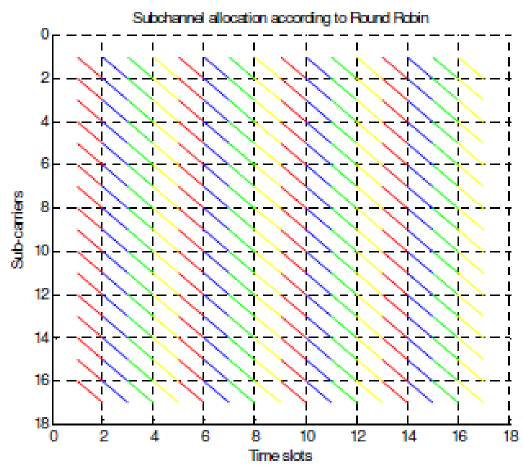
As we can see in Fig.4, the user A (red) is nearer to the BS than the other users because the channel A is stronger than others. In this scenario RR algorithm (see Fig:4 (a)) takes the same scheduler scheme that the previous scenario, it allocates system resources to one user at each time slot. When users have not the same channel response average, we can notice the differences between Max Rate and PFS on scheduling resources and the different PFS scheduler scheme when c t parameter changes. As we can see in Fig:4. (b), almost all subcarrier are allocated to user A because of channel strength and the same results are achieved with PFS policy when c t parameter is high (see Fig: 4 (c)). The larger c t parameter, the unfairer PFS is. RCG policy maintenances the index fairness because RCG allocates sub-carriers based on rate requirements and the users' average SNR.



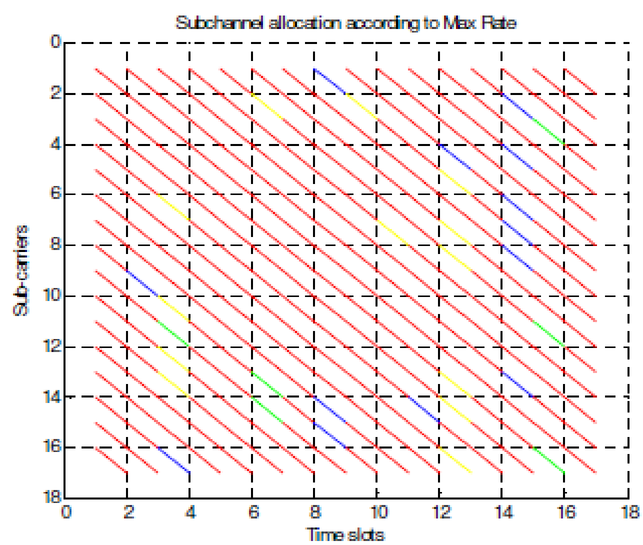
(a)



(b)



(c)



(d)

Fig. 4 Scheme where users have un equal channel responses average and have been simulated according to filtered Gaussian noise method, (a) RR algorithm, (b) Max Rate algorithm, (c) PFS algorithm, with tc =18, (d) RCG algorithm.

V. CONCLUSION

-MIMO is a promising technique which allows more than one user that can be served in each sub-carrier. Optimum number of users can be scheduled in a data region to achieve maximum sum capacity, which is equal to minimum number of antennas at the base station and the mobile station. The sum capacity increases linearly with SN R (dB) when number of users paired is not greater than minimum of the number of antennas at the BS and MS. In addition, we observe that as the number of user's increases, the sum capacity decreases.

REFERENCES

[1] L.Zheng and D. Tse " Diversity and multiplexing : A fundamental tradeoff multiple antenna chennels", IEEE Transactions on information theory, vol 49,no5 pp 1073-1096,May 2003. http://dx.doi.org/10.1109/TIT.2003.810646

[2] C. Gueguen and S. Baey, "Scheduling in OFDM wireless networks without tradeoff between fairness and throughput," in *Proceedings of the 68th Semi-Annual IEEE Vehicular Technology (VTC '08)*, September 2008.

[3] S.T.Chung,A.lozano,H.C.Huang, A.Sutivong and J.M. cioffi ," Approching the MIMO capacity with a low –rate feedback channel in V-BLAST", Eurasip Journal on applied signal processing vol51,pp 762-771,2004. http://dx.doi.org/10.1155/S1110865704312035

[4] R Van Nee and R.Prasad,"OFDM for wireless multimedia communications" Artech house, 2000.

[5] A.J.Paulraj,D.A.Gore,R.V.Nabar, and H.Bolcskei, " An overview of MIMO Communications a key to gigabit wireless ",Proceedings of the IEEE ,Vol 92,no2,PP 198-218,2004.

[6] Riberio C. B, Hugl K, Lampinen M, "Performance of Linear Multi-User MIMO Precoding in LTE system," *Third International symposium on Wireless pervasive Computing (ISWPC 2008)*, pp. 410-414, 2008.

[7] D. Kivanc, G. Li, H. Lui, *"Computationally Efficient Bandwidth Allocation and Power Control for OFDMA"* IEEE transactions on wireless communications, vol. 2, no. 6, November 2003. http://dx.doi.org/10.1109/TWC.2003.819016

[8] E. L. Hahne, *"Round-Robin Scheduling and Window Flow Control for Max-Min Fairness in Data Networks"* IEEE, technical report, November 1987.

Krishna.P: Completed B.Tech & M.Tech from JNTU-Hyderabad in 2005 and 2008.Pursuing Ph.D under the faculty of ECE with the specialization of LTE. His Research interested areas are wireless Communications & Networking, MIMO Communications, Multiuser-MIMO, Signals & image processing and a Member of IEEE

K.Kishan Rao: Presently working as Director in Vaagdevi group of institutions affiliated to JNTU-Hyderabad. He worked as Director and Principal in National Institute of Technology- Warangal. Under his supervision 12 Ph.D students are working in the area of Signals & Image Processing, and Wireless Communications& Networking. He is senior life Member of IEEE, ISTE.

Anil Kumar.T: Professor in the Department of Electronics & Communication Engineering of SR Engineering College- Warangal. Under his supervision more than 10 Ph.D students are working in the area of Signals & Image Processing, and Wireless Communications& Networking. He is senior life Member of IEEE, ISTE.