ABSTRACT
The next generations of wireless networks use OFDMA (Orthogonal Frequency Division Multiple Access) as the downlink radio transmission technology. With OFDMA, there is negligible interference between transmissions within a sector. However since transmissions in neighboring sectors may use the same frequency resources then such transmissions can cause interference. If this interference is too high then it negatively impacts the transmission rates achievable by users near the edge of the cell. Several approaches have been proposed to manage this cross-sector interference to achieve high sector throughput as well as acceptable performance for users near the cell edge. In this paper we summarize past approaches and then propose a simple distributed algorithm that requires no cross-sector coordination but achieves the desired performance characteristics. Since even the centralized version of this problem is intractable, our distributed algorithm is heuristic in nature but based on intuition gained from analytic results. Simplified simulation results are provided to support our claims.

Categories and Subject Descriptors
C.2.1 [Computer-Communication Networks]: Network Architecture and Design

Keywords
Interference Coordination, Self Optimizing Networks, 4G

1. INTRODUCTION
OFDMA is the chosen downlink radio transmission technology for the next generation of mobile communication systems (e.g., IEEE 802.16 [2] and LTE (Long Term Evolution) [1]. OFDMA offers flexible bandwidth support, high spectral efficiency and Multi-Input Multi-Output (MIMO) support. Intra-cell interference is avoided by scheduling at most one user in each time-frequency resource block (RB). However if a frequency reuse factor of one is used then transmissions in adjacent sectors may cause significant interference. If a frequency reuse factor of three is instead used then inter-cell interference can be reduced but this comes at the cost of capacity loss because of the reduced number of RBs available at each sector.

In this paper we present a simple distributed algorithm that each sector uses to determine an appropriate reuse factor. The algorithm aims to achieve two goals (a) acceptable interference levels at the edge of each cell and (b) fair resource allocation among sectors. Recall that the main objective of interference management is to be able to adequately service users near the cell edge. However we also need to maintain cross-sector fairness so that a sector’s throughput is not sacrificed while maintaining low cell edge interference levels at neighboring sectors. Note that in this paper we do not consider Multi-User MIMO whereby multiple users are served within a single RB.

In the literature there are two basic approaches to this problem, Soft Frequency Reuse (SFR) and Partial Frequency Reuse (PFR). In the SFR approach a subset of the frequency band is reserved for serving cell edge users. Typically this is one third of the total bandwidth. This subset is allocated with a frequency reuse factor of three so that three adjacent sectors use different parts of the bandwidth. The remaining bandwidth in each sector is used by the center users. Note however that the frequency partition used by edge users in one sector can also be used by cell center users in an adjacent sector. The edge users are served with higher power than the center users but the total transmitted power is maintained at a fixed level. If RBs reserved for cell edge users are not needed then they can be allocated to a cell center user. Note that because of the fixed partition of resources this approach cannot easily adapt to the wide variety of user loads, locations and QoS needs. In the PFR approach a subset of the bandwidth is also reserved for the edge users. This subset is further divided into three with each of the three portions being assigned to adjacent sectors so that a frequency reuse of three is achieved for the edge users. The cell center users are allowed to use the remaining bandwidth in each sector. Again note that this approach is not easily adaptable to the many possible variations of user loads, locations and QoS needs. These two approaches were compared in [3]. They concluded that the SFR approach provided the best trade-off between sector throughput and cell edge performance. However they did not investigate the case of inhomogeneous sector loading.

The thesis by Koutisimanis [4] considers the problem of joint channel and power allocation for the multi-cell OFDM
network. However the heuristic solution provided requires network-wide coordination whereas we focus on a distributed approach with no basestation coordination. The paper by Bosisio and Spagnolini [5] evaluates the performance benefit of interference coordination compared to interference randomization which requires no coordination. They conclude that, under light to moderate loading, interference coordination (using the PFR approach) is preferable but under heavy loading it is best to perform randomization. The Master’s thesis by Reider [6] proposes an algorithm in which users are partitioned into cell edge and cell center groups. Each group is served within a subset of resource blocks with the intent of allocating cell center users in resource blocks that can be assigned with a reuse factor of one while the cell edge users are allocated to the remaining blocks and these are assigned with a frequency reuse factor of three. The partition sizes are adjusted based on the user loading in each sector. However, under very heavy loading all resource blocks are used in each sector and one can show that such an approach is suboptimal. The thesis by Heyman [7] also contains an algorithm but it is static in nature and cannot adapt to changing network conditions.

In Section 2 we provide analytic results for some simple models to gain intuition of the problem solution. In Section 3 we present the proposed algorithm and in Section 4 we provide simulation results.

2. ANALYSIS OF A SIMPLE MODEL

In this section we consider analytic results for simplified versions of the problem and use the gained intuition to support the decisions made for our proposed algorithm. Note that similar results have been made in other papers in which more precise models are used. Our intent here is to show the intuition behind these results.

2.1 Bandwidth Utilization vs System Loading

We first demonstrate the fact that, although under light loading it may be optimal for all sectors to use a frequency reuse factor of 1, this is not necessarily the case for a heavily loaded network. We consider a simplistic model to illustrate this dependence.

Assume that each sector can use at most a total bandwidth \( B \) and that a single user is scheduled in each sector. This bandwidth can be divided equally among adjacent sectors to reduce interference. Assume that the bandwidth is divided equally among \( n \) sectors so that each sector experiences no interference from its closest \( n−1 \) neighbors. However it is still interfered by the other sectors. We estimate the interference as follows. Let \( A \) denote the area of a sector so that the radius \( r \) of the area covering the non-interfering sectors is given from the relationship \( An = \pi r^2 \). We assume that the sectors using the same band within a distance of \( \varepsilon \) outside this region interferes with the concerned sector. Furthermore, the interference falls inversely with the distance to the power of 3. We approximate the total interference as the product of the number of sectors in the surrounding ring \( 2\pi r\varepsilon/(An) \) times the loading \( \rho \) for the sub-band (i.e., the probability that the sector allocated to the sub-band actually uses it) times the interference caused by each \( nP/r^{-3.5} \) where \( P \) is the transmission power. We denote the background noise by \( N_0 \) and the path gain for the user served in the concerned sector by \( g \). Note that if the power density available for the whole band is \( P \) then for the sub-band it is \( nP \). We can now determine the Shannon capacity for the served user as

\[
R = \frac{B}{n} \log_2 \left( 1 + \frac{nP\rho g}{N_0 + 2n\rho r^{-3.5}} \right)
\]  
(1)

We now use the fact that \( r = \sqrt{An}/\pi \) in this equation and normalize the rate with respect to bandwidth to obtain

\[
R = \frac{1}{n} \log_2 \left( 1 + \frac{an}{1 + \rho bn^{-1.25}} \right)
\]  
(2)

where the parameters \( a \) and \( b \) can be determined from the constants \( P, g, A, N_0, \varepsilon \) and \( \pi \).

We determine reasonable values for \( a \) and \( b \) as follows. We assume that the interference becomes comparable to the background noise if we use a re-use factor of \( n = 6 \). Therefore \( b \approx 6^{1.25} \approx 10 \) and so we simply use \( b = 10 \). We also assume that if the entire bandwidth is used for the mobile, the SNR achieved by the mobile (i.e., without interference) is 10dB. This results in \( a = 10 \). We use these values together with a loading factor of \( \rho = 1 \) as the baseline problem.

We first investigate the affect that loading has on the optimal reuse factor. In Figure 1 we plot the baseline problem as well as the case in which we reduce the loading factor to \( \rho = 0.5 \). The horizontal axis is the FFR factor while the vertical access is the spectral efficiency of the transmission to the user in the concerned sector. First note that the spectral efficiency for the transmission to the user is higher for the lighter loading case (less interference). Secondly note that the optimal fractional frequency reuse factor moves from about 2 for the baseline problem to about 1.4 for the reduced loading case. In general higher loading is better supported with higher FFR which in turn means lower bandwidth utilization. The opposite also is true and for very lightly loaded cases each sector can use all available bandwidth.

![Figure 1: Effect of Loading on FFR Factor](image-url)

Next we investigate the effect that the position of the user in the sector affects the optimal FFR factor. In Figure 2 we again plot the baseline problem and we also include the case where the user’s SNR is 3dB lower (closer to the cell edge) than the baseline case. Note that the optimal FFR factor increases as the mobile approaches the cell edge. This fact, that users near the center should use a FFR factor of one while those at the edge should use higher FFR factors is well known.
that the interference level for a RB is the same for all users by the sector and that users achieved over $b_i$ is at least as high as when it was allocated to $j$. Hence the overall throughput can be increased (or at least remains the same) if $b$ is also allocated to $i$. Hence we can conclude that throughput can be maximized by assigning the user with the higher SINR to both RBs.

Note that, although user $i$ has the same channel gains over both RBs, the interference levels may be different and hence the SINR values may be different. This means that, in order to maximize throughput, it may be necessary to allocate different powers to each RB. Assume that the total power available for both RBs is $P$ and denote the individual powers by $p_{ia}$ and $p_{ib}$. Note that the rate achieved over a block increases with the allocated power. This means that the power constraint is binding and hence $p_{ia} + p_{ib} = P$. We simplify notation by denoting $p_{ia} = p_i$ and $p_{ib} = P - p_i$. We again consider the two sector case and denote the user allocated over the two blocks in the neighboring sector by $k$. The channel gains for this user, which is the same for each block, are denoted by $g_k$. The cross-sector channel gain from the concerned sector to the user in the neighboring sector will be denoted by $h_{ki}$. Similarly define $h_{ki}$. The powers allocated in the neighboring sector are denoted $p_k$ for resource block $a$ and $P - p_k$ for resource block $b$. The background noise of block $a$ in the concerned sector is assumed to be the same as for block $b$ and will be denoted by $N_i$. In the interfering sector the background noise is denoted by $N_k$. Note that in a homogeneous, well balanced network the interference from the other sectors in the network will be the same over each block. If this is the case then this component can be included in the background noise. Therefore, although we are only considering two sectors we can account for the other sectors as well (under the balanced loading assumption).

We can now determine the total spectral efficiency achieved in the concerned sector as

$$R_i = \log_2 \left( 1 + \frac{p_i g_i}{N_i + p_k h_{ki}} \right) + \log_2 \left( 1 + \frac{(P - p_i) g_i}{N_i + (P - p_k) h_{ki}} \right)$$

(3)

The power allocation that maximizes this rate is obtained from water-filling (essentially equating the rate derivatives with respect to power). If we do this we obtain,

$$p_i = \frac{P}{2} + \frac{(P - 2p_k) h_{ki}}{2g_i}.$$  

(4)

Since the neighboring sector also maximizes its throughput then it will also compute the optimal power allocation. For that sector we obtain

$$p_k = \frac{P}{2} + \frac{(P - 2p_i) h_{ki}}{2g_k}.$$  

(5)

Substituting for $p_k$ in the equation for $p_i$ and solving for $p_i$ we finally obtain $p_i = P/2$. In other words the available power is spread equally between the two resource blocks. We can repeat this for any two allocated resource blocks to conclude that one equilibrium point is uniform power allocation.

However this equilibrium point (Nash Equilibrium) is not necessarily optimal since it might be better for one sector to allocate all power to one RB while the other allocates all power to the other. Note that there are two possibilities only one of which will typically be optimal. By using power control the steady state solution has this property. However this limiting solution is equally likely to be either of the two options. We can see this as follows. Suppose that sector 1 has a high channel gain for RB $a$ and a low channel gain for RB $b$. Assume that the opposite is true for sector 2. The optimal solution would be for sector 1 to only use RB $a$ while sector 2 only uses channel $b$. Suppose that the system is in

![Figure 2: Effect of User’s Position on FFR Factor](image)
2.3 Fairness Issues

In the two previous subsections we focused on maximizing throughput. However, we must also take into account intra-sector fairness (fairness among users scheduled within a sector) and inter-sector fairness (fairness among users from different sectors). Intra-sector fairness is handled by the base-station scheduler and hence we focus on inter-sector fairness. First we must define what we mean by inter-sector fairness.

One definition of fairness is that all sectors must be allowed to use the same amount of time/frequency/power resources. For example, if interference is high then all sectors should use at most 80% of their bandwidth in order to limit interference levels. One problem with this approach is that for non-homogeneous user distributions resources may be wasted. For example if one region has a high user density while another has a low user density then the frequency reuse factor should be higher in the higher density region than in the lower density region. This means that a smaller fraction of the available bandwidth should be allocated to those sectors in the high density region than those in the low density region. By doing this fairness is maintained in that the allocation of more resources to users in the low density region does not adversely affect the performance of those users in the high density region.

Another definition of fairness is the following. Recall that the intent of interference coordination is to be able to limit the interference experienced by cell edge users thus allowing them to be able to achieve acceptable rates. Therefore we can define fairness as the allocation of resources to sectors such that the maximum interference experienced by any mobile is at most some specified value. This specified value will determine the trade-off between fairness and sector throughput. If this limit is very high then each sector will use almost all available resources and achieve high sector throughputs but the cell edge users will perform poorly. If the limit is low then few bandwidth resources will be used in each sector thus limiting overall sector throughput but the reduced interference will help the performance of the edge users. A suitable limit is determined by the type of application that must be supported and the associated QoS guarantees that must be provided no matter where the mobile lies within the cell. For example, for Voice over IP (VoIP) users the maximum interference level will be determined by the outage criterion.

We will take into account both of the above fairness criteria as follows. First we use the second fairness criterion and maintain an upper limit on the interference levels of the edge users. This is accomplished by having each sector independently change the fraction of bandwidth that it uses for transmissions. Since this results in different sectors using different bandwidth fractions then we also place a lower limit on this fraction. If this lower limit on the fraction of bandwidth used is very small then the interference criterion is more easily satisfied but the bandwidth used (and hence throughput achieved) by different sectors may differ drastically. If on the other hand the limit is large then all sectors have comparable resource allocations but the maximum interference limit criterion may be violated.

Note that in the above discussion we assumed that power is equally divided among all resource blocks and bandwidth was used to adjust interference levels. One can instead use all bandwidth resources and adjust power to adjust interference levels. More power can be applied for edge users than center users but this means partitioning of users. The partition used by one sector influences the interference levels experienced by its neighbors. This requires coordination among sectors. In our proposed approach there is no need for coordination among sectors.

3. PROPOSED ALGORITHM

In this section we present our coordinated interference management algorithm. From the results of the previous section we need to achieve the following:

- As the loading increases within a region of the network each sector in the region should reduce the number of resource blocks it uses for serving its users. The opposite is true for lightly loaded regions.
- The scheduler should continue to use the criteria for serving users that are used if no interference management is performed. In particular, frequency selective scheduling should not be affected. Note that this rules out approaches whereby the subcarriers of a resource block are spread over the entire bandwidth (interference randomization) since this prevents the use of frequency selective scheduling.
- Mobiles near the cell edge should achieve higher FFR factors than those close to the center.

We assume the following model (roughly based on the Long Term Evolution (LTE) standard). We assume \( N = 48 \) resource blocks within each subframe and that at most one user can be allocated to a resource block but that multiple of them can be allocated to a mobile. Each resource block consists of consecutive subcarriers and spans all symbols of the subframe. A collection of \( S = 4 \) consecutive resource blocks form a subband for a total of \( K = N/S = 12 \) subbands. Channel Quality Information (CQI) reporting and scheduling remains the same. However each mobile also periodically determines the intercell interference level experienced in each subband. This is also reported periodically.
(but at a much lower reporting rate than CQI reports). For each subband $b$, a sector takes the average over all users of the reported interference levels for that subband. Consider a particular sector and denote this average interference level by $\gamma_b$. This reflects the loading of this band. Based on the required performance for cell edge users we specify an interference threshold $T_{if}$. For each sector if $\gamma_b > T_{if}$ then band $b$ is not allocated for transmissions otherwise it can be allocated. In order to prevent specific sectors from blocking too many subbands (and hence sacrificing sector throughput) we place a lower bound $T_{sb}$ on the number of bands that can be blocked. If the number of subbands with average interference levels greater than $T_{if}$ exceeds $T_{sb}$ then only the bands with the $T_{sb}$ highest interference levels are blocked. We provide pseudo-code for the algorithm executed by each sector in Figure 3.

$$N = \text{number of users};$$
$$B = \text{number of subbands};$$
$$s_b = 1 \text{ for } b = 1 : B;$$
$$T_{if} = \text{interference threshold};$$
$$T_{sb} = \text{subband threshold};$$
$$\gamma_b = \text{Interference of mobile } i \text{ in subband } b;$$
$$\gamma_b = \frac{1}{N} \sum_{i=1}^{N} \gamma_{ib};$$
$$\text{for } k = 1 : T_{sb};$$
$$\text{if } (\max_b(\gamma_b) > T_{if}) \{$$
$$m = \arg \max_b(\gamma_b);$$
$$s_m = 0;$$
$$\gamma_m = 0;$$
$$\}$$
$$\text{end; Schedule users over subbands } b \text{ for which } s_b = 0;$$

Figure 3: Pseudo-code for proposed algorithm

Since we are only interested in the relative interference loading of the subbands to determine the ones that should not be used then, if interference levels are not reported we can instead use the reported channel quality information to estimate loading. First we assume that the channel quality indicator can be mapped to a SINR value. Second we assume that the variation of the channel gains across bands is much less than the variation in the interference levels across bands since the interference level changes each time a neighboring sector changes the set of RBs it uses for scheduling users. We also assume that the noise level is the same in each subband. Under these conditions the SINR value of a subband decreases with the interference level of the band. Hence for each user, the subband with the smallest SINR value is the one with the largest interference level. Hence for each subband we can take the average of the reported SINR values for the subband. We then order the subbands from smallest to largest SINR and use this as the loading order. Naturally this is an approximation and the actual interference reports are preferred.

Given the set of usable subbands, each sector then makes scheduling decisions in the same manner that it does when all subbands are usable. The subbands used by users near the cell edge will typically achieve a frequency reuse factor greater than one while those near the center will achieve factors close to one. This can be explained as follows. Assume that mobile scheduling priorities are proportional to their channel gains (e.g., this is the case for a Proportional Fair scheduler in which case the scheduling priority is proportional to the channel gain and inversely proportional to the user’s throughput). Consider the case of only cell edge users with high interference levels. Because of the interference threshold used in determining usable bands only a small number of subbands will be used to serve these users leading to a small bandwidth utilization (large reuse factor). On the other hand the opposite is true if all users are near the cell center. In the case of a mix of center and edge users the number of usable subbands depends on the distribution of the users because the interference loading is averaged across them. Therefore the bandwidth utilization will depend on the distribution of the user locations as well as the user loading in adjacent sectors.

4. SIMULATION RESULTS

In this section we illustrate the performance of the proposed distributed interference management algorithm through some simple simulation examples. We need to illustrate that (a) the fraction of bandwidth allocated to a sector is load dependent and (b) that the scheduler implicitly allocates mobiles to subbands according to their locations (center or edge) within the cell. An intuitive explanation of the latter property was given in the previous section but simulation results will be left for further study.

We consider a simple grid layout of square cells. Each base station lies in the center of a square and has neighbors in the adjacent squares. However cells at the edge of the region have fewer neighbors. Note that typically a wrap-around process is used so that all cells have the same number of neighbors. However we intentionally do not simulate wrap-around to illustrate how the algorithm adapts to non-homogeneous loading since those cells at the edge of the network have less interfering neighbors than those in the center. We determine path loss (inversely proportional to distance to the power of 3.5) from each base station to each mobile user and compute the received signal strength as well as the interference. Each cell independently updates its set of active subbands by computing the average interference for each subband and comparing with the threshold.

In our baseline problem we have $N = 30$ users per cell, $K = 12$ subbands, and a total of 25 cells. The background noise level was chosen to achieve a spectral efficiency of approximately 1 bps/Hz if a user is at the cell edge and there is no interference. For the baseline case a maximum of eight subbands can be blocked from use by a sector. The interference threshold used to determine whether or not a subband is blocked is normalized with respect to the maximum interference (over all users in all sectors) for the case in which all sectors transmit over all subbands. The default value of the threshold is set at one. We assume a simple round robin scheduler and hence the sector throughput is the average of the achievable mobile throughputs.

We are interested in the following three performance metrics (a) The average sector throughput (this will be normalized by the sector throughput for the case where all sectors use all subbands), (b) the maximum interference level over all users over all sectors (normalized by the same metric for the case in which all sectors use all subbands) and (c) the bandwidth utilization (this is the ratio of the number of subbands that a sector is allowed to use divided by the
The average sector throughput indicates the overall system level performance. The second metric provides an indication of the outage that will be experienced for those applications (like VoIP) for which a specified goodput must be achieved for acceptable user level performance. The bandwidth utilization metric indicates how the algorithm is achieving increased cell edge performance by dynamically changing the frequency reuse factors of cells based on the loading due to surrounding cells.

We first investigate the dependence of performance on the interference threshold. In Figure 4 we plot the three concerned metrics as a function of the interference threshold. Note that the interference threshold is normalized by the maximum interference over all users when all sectors use all subbands. First note that when the threshold is zero then all sectors will block subbands up to the maximum allowed. For the baseline case this is 8 subbands out of 12 and hence the bandwidth utilization is 1/3 or a frequency reuse factor of 3. When the threshold is 2 (high) then none of the sectors have blocked subbands and hence the bandwidth utilization is one. As the threshold increases from 0 to 2 the average sector throughput increases (good for overall system performance) but the interference level of the worst cell edge user also increases (increased outage). Note that as the frequency reuse factor increases from 1 to 3 the cell edge interference drops by approximately 20 percent. This translates into a significant improvement for cell edge users. This comes at the cost of a 43% reduction in sector throughput.

In Figure 5 we keep the interference threshold fixed at 1 and instead vary the maximum number of subbands that can be blocked by a sector. In this case the horizontal axis contains the maximum number of subbands that are allowed to be blocked. As this number is increased we find that the performance of cell edge users increases (less interference), the average sector throughput decreases and the bandwidth utilization droops (but is not equal to the maximum value).

We now keep the interference and subband thresholds fixed and instead vary the number of users per cell. In Figure 6 we provide the resulting performance plots. First note that the cell edge interference performance essentially flattens as the number of users increases. As the user population increases, the worst case user is statistically further away from its serving sector. Therefore it becomes more difficult to maintain that user’s throughput because of the increase in interference and the decrease in the channel gain from its serving sector. The algorithm is still able to maintain the cell edge performance. However this comes at the cost of reduced sector throughput and is achieved through increased frequency reuse factors.

In Figure 7 we plot the number of active subbands for each cell. Notice that the cells at all four corners of the network all use the full set of subbands (frequency reuse factor of 1) and this is reasonable since those cells are lightly loaded. The opposite is true for the cells in the center. These must block a subset of their subbands in order to provide acceptable interference to cell edge users in neighboring cells.

In Figure 8 we provide plots for each of the 12 subbands. Each plot contains 25 squares representing the 25 cells. If a cell is white then that subband has been turned off in that particular cell otherwise it is on. Here we see that for each subband the algorithm isolates a subset of cells so that
these cells can achieve acceptable interference levels at their edges. This is done in such a way so that different sectors are isolated in different subbands so that each sector can achieve similar sector throughputs. However the sectors on the edge of the network can achieve even higher throughputs because they have fewer interferers and hence can use more bandwidth.

For example, consider the two subbands in the third column. Look at the cells corresponding to columns three and four of these subbands. We clearly see the alternating on/off patterns for these cells. In traditional FFR approaches this pattern is enforced by coordination among the cells. However here it is achieved without coordination. Furthermore note that the cells at the edge have almost all subbands switched on so the algorithm adapts to non-homogeneous loading. This is difficult to achieve in traditional approaches.

5. SUMMARY AND CONCLUSIONS

In this paper we investigated the problem of managing the intercell interference for the downlink of an OFDMA network. We provided some intuition into the problem and then used this intuition to devise our proposed algorithm. The algorithm is self-adaptive (uses different frequency reuse factors for different sectors based on the conditions of their surroundings), is fair (the amount of bandwidth that each sector sacrifices is limited), and requires no intercell coordination (each sector infers adjacent sector loading information through measurements reported by its users). Simple simulation results were used to illustrate the performance of the approach. Future work will include more detailed simulation scenarios.

6. REFERENCES