This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the WCNC 2009 proceedings.

Distributed Load-Balancing in a Multi-Carrier Wireless System

Prashanth Hande, Shailesh Patil and Hyung G. Myung
Qualcomm Flarion Technologies, NJ 08807, USA

***Abstract*—We consider a cellular network or a wireless local area network (WLAN), deployed with attachment points (APs) capable of transmission and reception in multiple radio-frequency (RF) carriers. A major factor contributing to the efficiency and stability of the network is the mechanism determining the connection of terminals to the appropriate AP and the RF carrier. This paper describes a practical, distributed, network-assisted and terminal-driven mechanism to determine the connections for load-balancing among the available RF carriers. The mechanism is based on a metric that we term as the service level indicating metric (SLIM). We extend the notion of SLIM to cases where Quality-of-Service (QoS) parameters are specified and propose a load-balancing mechanism for such cases. We demonstrate the near-optimality of the proposed mechanisms through 3GPP based simulations.**

Keywords: Base Station assignment, load-balance, multi-carrier, handover, distributed mechanism.

I. Introduction

Wireless networks have experienced phenomenal growth in the last twenty years and the trend is expected to continue further with wider adoption of 3G, 4G and wireless LAN networks. The capacity or the data throughput share of individual users is reduced with increasing number of users in such networks. The primary mechanism to increase capacity is addition of further spectrum to the network[1]. Wireless networks are, however, designed to operate in a fixed bandwidth, referred to as the channel or the carrier. For example, the EVDO-RevA system was designed to operate in a bandwidth of 1.25MHz and the 802.11b based wireless LAN system was designed to operate in a bandwidth of 22MHz. Although next generation wireless systems are typically being designed to provide flexibility in terms of the bandwidth of operation, it might be unreasonable to wait for the availability of these systems to improve the capacity of the network. In addition, if the newer systems do not support existing wireless devices, then the capacity upgrade might be an expensive operation. In light of this, a preferred approach to the capacity problem is to stack multiple radio-frequency carriers or channels as shown in Figure 1.

In such a multi-carrier network, the Attachment Points (AP), Base Stations in cellular networks and Access Points in WLAN, can connect to wireless terminals (WT) on one or more of several radio-frequency (RF) carriers. For example, a cellular network based on EVDO-RevB [1] is one such system where multiple carriers can be aggregated to enhance the capacity. Another example is a wireless LAN system deployed in infrastructure mode with each AP transmitting on multiple orthogonal wireless LAN channels [2], [3]. A third example is a network based on Flash-OFDM where the network can transmit on up to three 1.25Mhz wide carriers [4]. The total capacity available in the multi-carrier network scales with the number of carriers that are supported. The network can provide service to WTs that have *different radio-frequency capability*. For example, legacy WTs or inexpensive WTs, designed to keep the cost and power consumption low, can have single-carrier radio-frequency capability while WTs with higher data rate requirements can be capable of multiple carriers.

[1]Cell-splitting can result in capacity increase only up to a certain extent

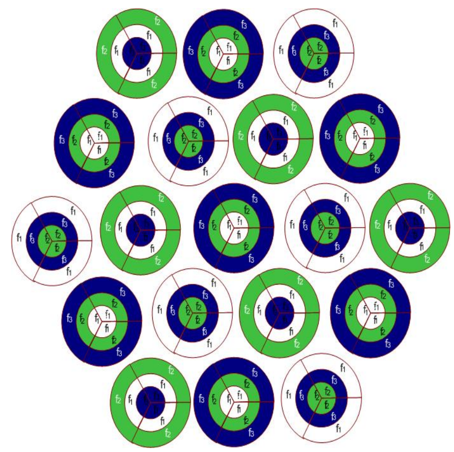

Fig. 1. Example of a multi-carrier system with $C = 3$ carriers with carriers denoted by f_1, f_2, f_3 in each cell. Carrier transmission powers can be different and is proportional to the radius of the circle representing that carrier in the figure.

The APs can transmit multiple carriers at equal power levels or at different power levels, with power levels designed to provide different frequency re-use factors. For example, the Flash-OFDM network implements a scheme called 'Flexband' [5], variations of which have been adopted in next generation wireless systems, and known variously as fractional power reuse, fractional frequency reuse and partial frequency reuse [6]. A simple SIR or signal power level based connection mechanism is insufficient with multiple carriers. The SIR and signal power measurements will be near-identical when multiple carriers on an AP are transmitted at equal power levels. With multiple carriers at different power levels, the carrier of the highest power level will be the preferred connection, resulting in a non-optimal configuration. The performance of a multi-carrier network depends on the *load-balancing* mechanism that determines the WT connections.

The EVDO-RevB network and many enterprise deployments of the Wireless LAN network employ a centralized controller or a switch that can decide on the appropriate connections. The centralized controller determines the connections of the WTs to the appropriate AP and the carrier, achieving the appropriate load-balancing. Load balancing in wireless LAN has been addressed primarily through centralized mechanisms in [7], [8]. The Flash-OFDM network, on the other hand, has an architecture that does not include a centralized controller, and the connection decision is made by the WT based on information gathered from the network. A "flat" network architecture [9] with no centralized controllers is increasingly the trend in future wireless system designs, and offers advantages in cost and simplicity of the network. This paper describes a *network-assisted, terminal driven*, distributed load-balancing mechanism that is appropriate for such networks.

The distributed mechanisms proposed in this paper has the simplicity required for a *practical real network* implementation, and yet achieves the near-optimal configuration desired. Variations of the proposed mechanisms have been implemented as part of the Flash-OFDM system. This paper introduces the set-up and presents the mechanisms, along with simulations that demonstrate the near-optimality. Analysis of optimality and convergence is expected to stem from future work. The major contributions of this paper are summarized below.

- We introduce a metric called the "Service Level Indicating Metric (SLIM)", a metric that provides a measure of the user experience and is the basis for our load-balancing mechanisms.
- For a network offering best effort service with no quality of service (QoS) differentiation, we propose a mechanism that achieves load-balance across multiple carriers in the network. We demonstrate through simulations that this mechanism achieves a performance close to optimal in terms of a network-wide throughput criterion.
- For a network offering QoS, we propose a QoS aware load-balancing mechanism, that achieves load-balancing between multiple carriers to minimize the outage probability of WTs with QoS requirements, while maximizing the network-wide throughput criterion for non-QoS WTs.

The paper is organized as follows. Section II describes the system setup and the load-balancing mechanism. The mechanism is extended to support quality-of-service (QoS)-requirements in Section III. Simulation methods and results are presented in Section IV and we conclude with Section V.

II. Load Balancing Mechanism

A. System Setup

Consider a system with N WTs and M APs where each AP transmits on C carriers as shown in Figure 1. For simplicity, let the terminals be all single-carrier capable. The terminals establish a connection to a given AP on one of the carriers transmitted by the AP. Let σ_i denote the AP and c_i the carrier on that AP that terminal i connects to. Let $\mathcal{A}(k, c)$ be the set of all WTs that are connected to cth carrier transmitted by AP k. Let $L(k, c)$ be a number which we call the carrier load-factor corresponding to AP k and carrier c, representing the usage of capacity on that carrier. For now, we set $L(k, c) = \|\mathcal{A}(k, c)\|$, where $\|.\|$ denotes the cardinality of the set so that $\|\mathcal{A}(k, c)\|$ is the number of WTs sharing that connection.

Load-balancing is facilitated by information that the terminal obtains from the network. It is assumed that each AP exchanges information with a set of neighboring APs over the back-haul connection. The mechanism that enables each AP to build a neighbor-list and communicate with those neighbors without the assistance of a centralized controller, is outside the scope of this paper. The communication between the APs enables each AP to accumulate local network information relevant for the load-balancing mechanism. This information is broadcast by the AP on each of the carrier it transmits on and is read by the terminals connected to the AP.

The broadcasted network information relevant to the load-balancing mechanism includes the carrier load-factor. In addition, if multiple carriers are transmitted at different power levels, then the relative carrier transmit-power levels are also included in the broadcast information. The WT measures the strength of the radio-frequency signals by measuring the received power of the acquisition signal from a set of APs in the vicinity of the terminal. This information, combined with information on the relative carrier power levels that the WT reads from the broadcast, enables the WT to derive the signal-to-interference ratio (SIR) that it would potentially experience on a connection.

B. Service Level Indicating Metric (SLIM)

A simple criterion for selecting a connection can be that the measured SIR for that connection. However, the SIR measurement results in an uncertain choice among carriers with equal powers, and biased choice with carriers at different power levels. So it is imperative that the connection selection decision take into consideration factors beyond the measured SIR to achieve load-balancing. The user experience depends upon the SIR and the allocation of time slots and frequency bins, referred to as time-frequency resources in the rest of the paper. We present a metric that captures the user experience under different schemes used for allocation of the time-frequency resources.

The load-factor information is combined with the SIR to calculate a metric that we refer to as the Service Level Indicating Metric (SLIM) which will be indicative of the user experience of the WT on a given connection.

$$s_i(k,c) = \begin{cases} \frac{\log(1+\gamma_i(k,c))}{L(k,c)} & \text{if } k=\sigma_i, c=c_i \\ \frac{\log(1+\gamma_i(k,c))}{1+L(k,c)} & \text{otherwise} \end{cases} \quad (1)$$

where $\gamma_i(k,c)$ is the SIR on AP k and carrier c calculated by the WT by obtaining appropriate information from the network. The SLIM value $s_i(k,c)$ is an indication of the data rate that the WT can obtain if it were to connect to AP k on carrier c. If the time-frequency resource allocation policy is that of equal allocation to all the connected terminals, then the data rate experienced by each WT will indeed be proportional to the SLIM associated with that connection.

In practice, the resources are not equally allocated for various reasons including provisioning QoS treatment to WTs. The SLIM definition can be modified to incorporate alternate time-frequency resource allocation policies. For example, a weight-based resource allocation policy does not grant any WT a strict priority over others, but allocates time-frequency resources to WTs according to weights, specified to reflect relative priorities between WTs. A network implementing this resource allocation policy can set the carrier load-factor to

$$L(k,c) = \sum_{i \in \mathcal{A}(k,c)} w_i \quad (2)$$

The SLIM calculation changes to

$$s_i(k,c) = \begin{cases} \frac{w_i \log(1+\gamma_i(k,c))}{L(k,c)} & \text{if } k=\sigma_i, c=c_i \\ \frac{w_i \log(1+\gamma_i(k,c))}{w_i+L(k,c)} & \text{otherwise} \end{cases} \quad (3)$$

Further extension of the SLIM metric to cases involving WTs with strict QoS requirement and a resource allocation policy that strictly prioritizes QoS WTs are discussed in Section III.

C. Load-Balancing Mechanism

We propose the following scheme for selecting the appropriate connection for each WT i.

Algorithm *1* **(Distributed Load Balancing Mechanism):**

At each WT i:

1) Calculate the SIR $\gamma_i(k,c)$ from the network information and a measurement of the acquisition signal.
2) For each carrier c, select AP $\tilde{k}(i,c)$ that results in the maximum SIR:

$$\tilde{k}(i,c) = \text{argmax}_k \gamma_i(k,c), \ c = 1,\ldots,C \quad (4)$$

3) Calculate the SLIM for each connection in the set $\{\tilde{k}(i,c), c = 1,\ldots,C\}$.
4) Pick the carrier c_i with the highest SLIM

$$c_i = \text{argmax}_c s_i(\tilde{k}(i,c),c) \quad (5)$$

and connect to AP $\tilde{k}(i,c_i)$.

Remarks:

1) In practice, the WT can calculate the SIR for only the connections on which the acquisition signal can be detected and measured by the WT.
2) The number of SLIM calculations required is equal to the number of carriers supported by the network.
3) The mechanism ensures that the WT is always connected to the AP of maximum SIR on a given carrier.
4) The mechanism ensures that the WT is connected to the carrier on which it measures the highest SLIM.

Notice that the mechanism chooses between multiple carriers based on SLIM but within a carrier the choice of the AP is based on the SIR. This ensures the stability of the cellular architecture whether it is a wide area network based wireless system or a wireless LAN system operating in the infrastructure mode.

III. QoS Aware Load-Balancing

We extend the SLIM based load-balancing mechanism to incorporate quality-of-service (QoS) framework. We consider a priority-based QoS policy. In a network implementing priority-based QoS, a set of WTs (QoS-WTs) $\mathcal{Q}(k,c)$ at each AP k with QoS requirements $\{\beta_i\}$ have strict priority over non-priority WTs (BE-WTs), who are offered a best-effort service and grouped into set $\mathcal{B}(k,c)$.

With QoS-WTs, it is no longer sufficient to design a load-balancing mechanism that optimizes a network-wide criterion based on WT data-rates. Instead, the mechanism has to take into account the likelihood of satisfying requirements of the QoS-WTs, while optimizing a network-wide criterion based on data-rates of BE-WTs. With strict prioritization of QoS traffic, the likelihood of a connection satisfying QoS requirement is a function of the number of QoS-WTs on that connection alone, and is independent of the number of BE-WTs. The data rate of a BE-WT is a function of both the number of BE-WTs and the number of QoS-WTs on that connection. We propose a QoS aware load-balancing mechanism that takes into account these considerations.

The resource allocation mechanism operating in the network is assumed to strictly prioritize QoS traffic by reserving a fraction of time-frequency resources. Let $F(k,c) \leq 1$ denote the fraction of time-frequency resources utilized by AP k on carrier c to serve QoS-WTs. We do not go into the precise resource allocation mechanism through which the QoS requirements are met but discuss possible ways to calculate the fraction $F(k,c)$ at the end of this section. The network information is now assumed to consist of a carrier load-factor $L(k,c) = \|\mathcal{B}(k,c)\|$ reflecting the number of BE-WTs on a connection and the fraction $F(k.c)$. The QoS WTs in set $\mathcal{Q}(k,c)$ calculate SLIM as

$$s_i^q(k,c) = (1 - F(k,c)) \log(1 + \gamma_i(k,c)) \quad (6)$$

The Best-Effort WTs in set $\mathcal{B}(k,c)$ calculate SLIM as

$$s_i^b(k,c) = \begin{cases} \frac{(1-F(k,c)) \log(1+\gamma_i(k,c))}{L(k,c)} & \text{if } k=\sigma_i, c=c_i \\ \frac{(1-F(k,c)) \log(1+\gamma_i(k,c))}{1+L(k,c)} & \text{otherwise} \end{cases} \quad (7)$$

The load-balancing mechanism remains the same as in Algorithm 1. The fraction $(1 - F(k,c))$ indicates the fraction of time-frequency resources that are available to existing BE-WTs and any QoS-WT that intends to establish connection. The SLIM value for BE-WTs $s_i^b(k,c)$ is an indication of the data rate that the BE-WT can obtain if it were to connect to AP k on carrier c. The SLIM value for QoS-WT $s_i^q(k,c)$ is an indication of the resources available to satisfy the QoS requirement β_i at AP k on carrier c.

The calculation of the fraction $F(k,c)$ of time-frequency resources used to serve the QoS-WTs will, in general, depend upon the implementation of the mechanism that allocates the time-frequency resources. For example, this mechanism is implemented through an explicit scheduling algorithm in Flash-OFDM and through a contention-based algorithm in wireless LAN. However, we can provide the following guideline for the calculation of this fraction. Suppose the QoS requirement β_i is a fixed rate r_i for the QoS user i. If B is the frequency bandwidth of the system, then a reasonable estimate of the fraction f_i of the time-frequency resource can be given by

$$f_i = \frac{r_i}{B \log(1 + \gamma_i(\sigma_i, c_i))} \tag{8}$$

If the QoS requirement β_i is a tuple consisting of fixed rate r_i and delay τ_i, and we can further make M/M/1 assumption on the packet queue model [10], then the fraction can be calculated as

$$f_i = \frac{r_i + 1/\tau_i}{B \log(1 + \gamma_i(\sigma_i, c_i))} \tag{9}$$

where we have assumed that the delay is predominantly queuing delay and the WT data rate with fraction allocation f_i is $f_i B \log(1 + \gamma_i(\sigma_i, c_i))$. In either case, the fraction $F(k,c)$ for each connection is given by

$$F(k,c) = \sum_{i \in \mathcal{Q}(k,c)} f_i \tag{10}$$

We note that the guideline used for the calculation of the time-frequency fraction to satisfy a given QoS requirement β_i, can also be used to as a guideline for distributed admission control[2]. The admission control mechanism determines the likelihood of satisfying the QoS requirements of a QoS-WT, and permits access to the WT if the likelihood is sufficiently high. With a fixed rate r_i as the QoS requirement of WT i, the following distributed admission control decision can be employed: Admit if $(1 - F(k,c)) \log(1 + \gamma_i(\sigma_i, c_i)) - r_i > 0$ and do not admit otherwise with σ_i and c_i being the AP and carrier respectively, selected through the load-balancing mechanism.

IV. Simulations

We demonstrate the performance of the SLIM-based load-balancing mechanisms in a cellular network, using a simplified version of the 3GPP network and path loss models [11] that are adopted by the wireless industry.

We can consider the performance of the load-balancing mechanism by defining an objective measure on the data rates $\{r_i\}$ allocated to the WTs. One such measure is the maximization of the sum of the logarithm of the data rates of all the WTs in the network[3] subject to feasibility condition:

$$\begin{array}{ll} \max & \sum_i \log(r_i) \\ \text{subject to} & \{r_i\} \in \mathcal{R} \end{array} \tag{11}$$

where $\mathcal{R}$ represents the feasible data rate region across all possible WT connections. The solution to the above problem requires a search over all possible combinations of WT connections to the network and is computationally expensive. To demonstrate our mechanism, we consider an approximation to the optimal scheme which we call the 'optimal approximation'(OA) mechanism. Let $\mathcal{A}(k)$ be the set of all WTs connected to AP k and let $r_i(k,c)$ be the WT data rate when connected to AP k and carrier c.

Algorithm *2* *(***Optimal Approximation (OA)***)***:**

---

1) Calculate the SIR $\gamma_i(k,c)$ for each i, k, c.
2) Select the AP for each i that results in the maximum SIR across all carriers:

$$\tilde{k}_i = \text{argmax}_{k,c} \gamma_i(k,c) \tag{12}$$

3) Pick $\tilde{k}_i$ as the AP for WT i and pick carriers within AP $\tilde{k}_i$ such that

$$c_i = \text{argmax}_c \sum_{j \in \mathcal{A}(\tilde{k}_i)} \log(r_j(\tilde{k}_i, c)) \tag{13}$$

In effect, the OA calculation first selects the AP for each WT based on a SIR maximization criterion, and then allocates WTs to carriers within each AP through a maximization of the sum of the logarithm of the rates. This is a reasonable approximation to the optimal since we simulate cases where the spread of WTs is even across APs. The focus then is on the performance of the load-balancing mechanism in determining the appropriate carrier within each AP so as to maximize the overall objective.

The basic assumptions and parameters for the simulations are in Appendix A. We assume that there are three carriers in each sector. We consider two different cases of carrier power levels within each AP.

1) Equal Power (EP) on all 3 carriers.
2) Fractional Power (FP) configuration where the strongest carrier is 6dB stronger and the weakest carrier is 6dB weaker than the third carrier.

Figures 2 and 3 plot the cumulative distribution (CDF) of the SIR and the data rates for the cases with equal power carriers and fractional power carriers. The distribution of both the SIR and the data rates with our load-balancing mechanism

[2]Admission control is the mechanism used in cellular networks to determine if a terminal should be allowed to connect to the network.

[3]Such a measure is termed *proportional fairness*.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the WCNC 2009 proceedings.

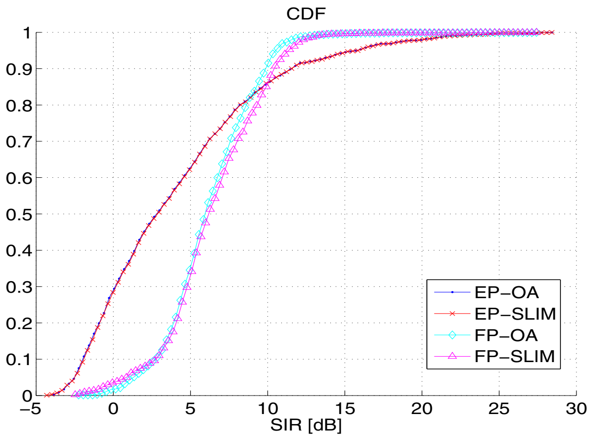

Fig. 2. Comparison of SIR distribution.

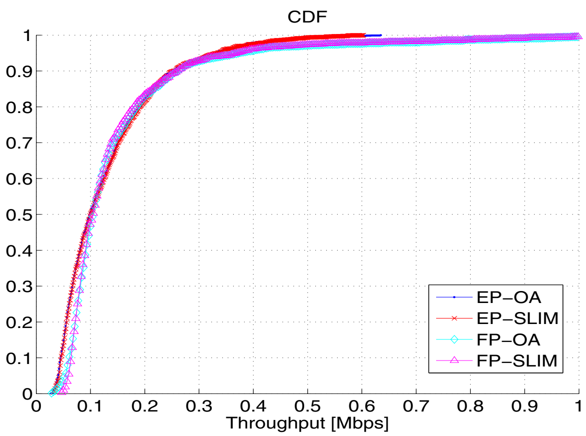

Fig. 3. Comparison of data rate distribution.

is similar to the optimal approximation. The performance gain in the fractional power case as shown in Table I is retained even with our distributed load-balancing mechanism.

In networks with QoS requirements, the load-balancing mechanism needs to strike a balance between satisfying the QoS requirements and achieving high data rates for the best-effort WTs. We assume that QoS WTs have a fixed rate requirement of $r_i = 100$kbps. Table II shows the performance of the QoS aware mechanism introduced in Section III when the average number of WTs in each sector is varied from 20 to 60, with equal number of QoS and Best-Effort WTs. In addition to throughput of BE-WTs, we tabulate the outage probability of QoS-WTs, which is the percentage of QoS-WTs that cannot be allocated an average of 100 kbps of fixed rate requirement over multiple instances of a shadow-fading channel. It can be noticed that QoS aware SLIM provides significant improvement in outage probability for QoS WTs. Figure 4 plots the CDF of the data rates of all WTs in the network.

TABLE I
ACHIEVED DATA RATES

	Equal Powers	Fractional Power
Avg. AP data rate	7549 kbps	9240 kbps
Avg. WT data rate	130 kbps	148 kbps
Median WT data rate	99 kbps	105 kbps
90 % WT data rate	250 kbps	105 kbps
10 % WT data rate	46 kbps	64 kbps

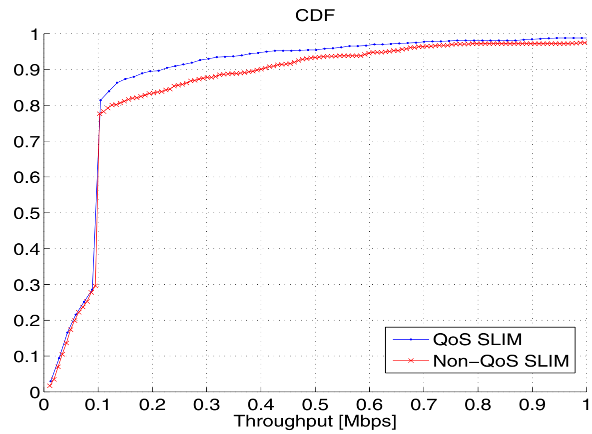

Fig. 4. Data rate distribution with 40 WTs per sector on average.

V. Conclusion

We proposed a distributed mechanism for load-balancing in a cellular-like wireless network with attachment points deploying multiple carriers. We introduced a metric, SLIM, based on which the distributed mechanism achieves load-balancing. The SLIM metric was extended to the case where WTs need QoS treatment. A load-balancing mechanism to achieve an appropriate trade-off between QoS outage and throughput was proposed. Simulations with realistic 3GPP based models demonstrate the near-optimality of our mechanism.

Although we demonstrate the near-optimality of our mechanism through simulations, it would be desirable to analyze the performance of the mechanism in a mathematical framework to demonstrate either the optimality or the distance to optimality of the mechanism. In addition, the convergence of the algorithm needs further analysis. We expect such analysis to stem from future work.

Load-balancing mechanisms operate with one degree of freedom, the choice of an appropriate connection for each WT, to achieve a network-wide objective. A wireless network provides additional degrees of freedom in the form of power, time-slots, bandwidth of transmission, and use of multiple antennas that provide a spatial degree of freedom. The appropriate use of these degrees of freedom to achieve a network-wide objective encompasses the functionality of a scheduling algorithm. Further work is required to design a joint scheduling and load-balancing mechanism to maximize a network-wide objective.

TABLE II
OUTAGE AND DATA RATES FOR QOS AWARE SLIM (Q-SLIM) AND NON-QOS SLIM (NQ-SLIM)

Number of Users	Best-Effort Rate (Q-SLIM)	QoS Outage (Q-SLIM)	Best-Effort Rate (NQ-SLIM)	QoS Outage (NQ-SLIM)
20	6.01Mbps	0%	8.08Mbps	0.01%
30	4.37Mbps	0%	6.25Mbps	1.18%
40	4.06Mbps	0.02%	5.09Mbps	5.51%
50	2.49Mbps	1.43%	4.74Mbps	13.37%
60	1.24Mbps	6.21%	4.05Mbps	20.53%

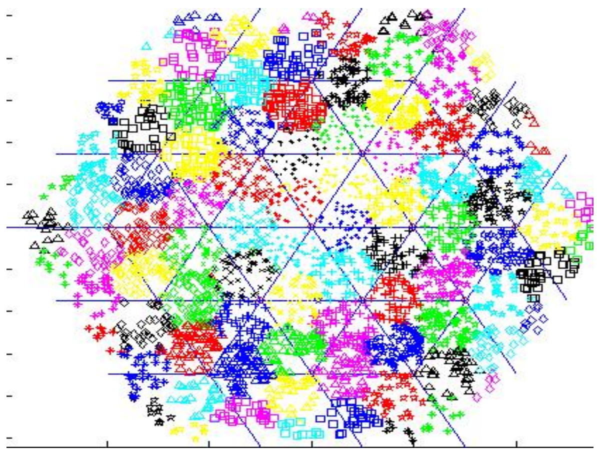

Fig. 5. Distribution of WTs across cells and their connections.

ACKNOWLEDGMENT

The authors acknowledge Pablo Anigstein, Xinzhou Wu and Arnab Das.

APPENDIX A: SIMULATION PARAMETERS

The following are the basic assumptions for the simulations.

A. Cell layout and configurations

1) Hexagonal grid 19-cell wrap around layout (Figure 1): Only the statistics of the users allocated to the inner ring are considered in order to minimize the imperfect wrap around configuration for the Fractional Power case.
2) 3 sectors per cell
3) Cell-to-cell distance: 1 km
4) Minimum distance between the mobile terminal and the cell site: 35 m

B. Antenna configurations

1) 3 dB cutoff angle: $\theta_{3dB} = 65^o$
2) Front-to-back loss: $A_m = 32$dB
3) Antenna pattern:

$$A(\theta) = -\min\{12\left(\frac{\theta}{\theta_{3dB}}\right)^2, A_m\} \text{ dB}$$

4) Antenna height AP: $h_{AP} = 32m$, WT: $h_{WT} = 1.5m$

C. Radio configurations

1) Carrier frequency: $f_c = 450$ MHz
2) BW per carrier: 1.271250 MHz
3) Number of carriers: $N = 3$ (Equal Powers or Fractional Power) In Fractional Power configuration, the power of the strongest carrier is 6 dB higher than that of the second strongest carrier and the power of the weakest carrier is 6 dB lower than that.

D. Propagation

1) Distance-dependent path loss:

$$\begin{aligned} PL(r) &= 46.3 + 33.9\log_{10}(f_c) - 13.82\log_{10}(h_{AP}) \\ &+ (44.9 - 6.55\log_{10}(h_{BS}))\log_{10}(r) \\ &- (1.1\log_{10}(f_c) - 0.7)h_{WT} \\ &+ (1.56\log_{10}(f_c) - 0.8) \end{aligned}$$

for $r$km, f_cMHz, h_{AP}m, h_{WT}m.
2) Shadowing introduced for QoS results.

E. Mobile terminal configurations

1) There are $20 - 60$ users per AP on average.
2) Mobile terminals are dropped within the radius of the cell and are uniformly distributed within each cell.

REFERENCES

[1] R. Attar, D. Ghosh, C. Lott, M. Fan, P. Black, R. Rezaiifar, P. Agashe, "Evolution of cdma2000 cellular networks: multicarrier EV-DO", IEEE Communications Magazine, March 2006, Volume: 44, Issue: 3, page(s): 46- 53

[2] Cisco White Paper, "The Benefits of Centralization in Wireless LANs", http://www.cisco.com

[3] Aruba White Paer, "WLAN RF Architecture Primer: Single-Channel and Adaptive Multi-Channel Models" http://www.arubanetworks.com

[4] R. Laroia, S. Uppala, J. Li, "Designing a mobile broadband wireless access network," IEEE Signal Processing Magazine, Volume 21, Issue 5, Sept. 2004 Page(s): 20 - 28

[5] X. Wu, A. Das, J. Li, and R. Laroia, "Fractional Power Reuse in Cellular Networks," *44th Annual Allerton Conference*, Sep. 2006.

[6] K. Son, S. Chong and G. de Veciana, "Dynamic Association for Load Balancing and Interference Avoidance in Multi-cell Networks," WiOpt 2007, Apr. 2007.

[7] H. Velayos, V. Aleo, G. Karlsson, "Load balancing in overlapping wireless LAN cells," *IEEE International Conference on Communications*, June 2004

[8] A. Balachandran, P. Bahl, G. M. Voelker, "Hot-Spot Congestion Relief in Public-area Wireless Networks," *Proc. of 4th IEEE Workshop on Mobile Computing Systems and Applications*, June 2002

[9] M. S. Corson, R. Laroia, A. O'Neill, V. Park, G. Tsirtsis, "A New Paradigm for IP-Based Cellular Networks," *IT Professional*, Nov-Dec 2001.

[10] D. Bertsekas, R. Gallager, *Data Networks*, Prentice Hall, 1999.

[11] 3GPP TR 25.896 (2004-02), *Feasibility Study for Enhanced Uplink for UTRA FDD*, Document R1-040346, V1.3.0.