In Proc. of the 16th Annual IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC 2005), Berlin, Germany, September 2005

Load Balancing for Centralized Wireless Networks

Hong Bong Kim and Adam Wolisz
Telecommunication Networks Group
Technische Universität Berlin Sekr FT5 Einsteinufer 25 10587 Berlin Germany
Email: {hbkim, wolisz}@tkn.tu-berlin.de
Telephone: +49-30-314-24576, Fax: +49-30-314-23818

Abstract— **Recently, centralized wireless network architecture has been proposed for efficient system resource management and cost-effectiveness. Radio over fiber (or cable) networks and centralized wireless LAN systems are representative instances. However, this architecture requires us to reconsider system level control schemes since most of research efforts in this field has been focused on distributed network architecture. In this paper, we propose a load balancing scheme for centralized wireless networks using the capability that users in the overlapping area between cells can efficiently and dynamically be reconnected to one of the accessible base stations by a central control station. A simulation study is done to compare the proposed scheme with conventional load balancing ones, showing the proposed scheme is more efficient in system resource utilization.**

I. Introduction

Recently, centralized wireless network architecture has begun to receive attention for its capability for efficient system resource management and cost-effectiveness. Representative examples include radio over fiber (RoF) based (or radio over cable) networks [1] and centralized architecture for IEEE 802.11x WLAN systems [2]. In RoF networks, a central control station (CS) is connected to multiple base stations (BSs) via optical fiber and BSs serve as access points for subscribers. One peculiar feature of RoF networks that is quite different from conventional wireless networks is that all signal processing including modulation, demodulation, medium access control and so on is done in the CS and BSs serve as simple optic-wireless converter and vice versa. As a result, RoF networks become inherently centralized networks. On the other hand, in a first-generation IEEE 802.11 wireless LAN, the network's intelligence is distributed among the access points (APs). But resource management techniques require access to information that must be gathered across a number of APs, and the techniques involve control decisions that apply to a number of APs, not just one. Thus, some centralized decision making is appropriate. In [2] a centralized architecture that makes this possible has been proposed, where an intelligent switch serving as a CS controls APs deployed throughput the target space.

Resource management issues in wireless networks based on cellular architecture involve dynamic channel assignment, dynamic transmit power control, load balancing, mobility management and so on. Since research efforts in this field has mainly been focused on distributed wireless architecture, we are required to *"reconsider"* system level control schemes using centralized control capability. In [3], [4] mobility management schemes using such capability of RoF networks have been proposed.

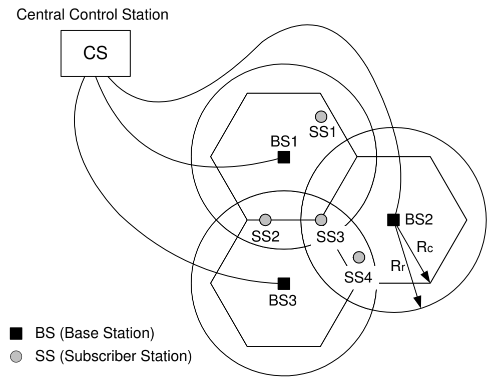

Fig. 1. Centralized wireless access network architecture. R_c is the cell radius and R_r is the radio coverage radius.

In this paper we propose a dynamic load balancing algorithm achieving efficient system resource management for centralized fixed broadband wireless access networks. Basic idea of load balancing scheme in conventional cellular networks is to use subscribers located in overlapping region between cells for better performance in terms of connection blocking probabilities and channel utilization [6]– [10]. However, due to its distributed architecture dynamic load balancing that can be made more efficient when central decision making is possible is not simple.

The paper is organized as follows. Section II gives a description of system and traffic model, admission control and the proposed load balancing algorithm. In section III we evaluate the algorithm by a simulation study. Finally, conclusions are drawn in section IV.

II. Model Description

A. System Layout

A cellular system with omnidirectional BS antennas is considered that are organized in a hexagonal pattern as shown in Fig. 1. The service area is partitioned according to the number of available BSs. In particular, A_i is defined to be the area where the number of available BSs is i. For instance, in Fig. 1 subscriber stations (SS) 1 is located in A_1, SS2 and SS4 are in A_2, and SS3 is in A_3. Cell radius (R_c) is defined as the distance between center of the cell and a vertex of hexagon. The coverage of a BS is the area in which an SS can establish a link with acceptable signal quality with the BS. This area is modeled by a circle with the BS at the center. The radio coverage radius (R_r) is defined as the distance from a BS to its coverage boundary. In this paper R_r represents a normalized value relative to R_c from now on. Once R_r is determined, the percentage of a cell that belongs to region $A_i, i = 1, 2, \ldots$ can be calculated using the equations provided in [5]. Each cell is assigned a fixed number of channels (C); that is, fixed channel assignment is assumed.

When an SS wants to associate with the network, it first makes a list of available BSs. After that the SS sends bandwidth requirement along with the list to the CS via the BS with highest quality signal. The CS maintains a table consisting of SS's identification (ID) number, required bandwidth and its accessible BS IDs. In this paper the procedure is assumed to be done when an SS associates with the network.

B. Traffic Model

Traffic arriving at each cell is partitioned into M separate classes based on bandwidth requirements. Each class i requires a bandwidth of B_i channels where B_i is an integer value that might be time slots in TDMA system or frequency channels in FDMA system. Let $\mathbf{B}$ denote a vector representing bandwidth requirements of each class $\mathbf{B} = (B_1, B_2, \ldots, B_M)$. Similarly, define subscriber vector $\mathbf{N} = (N_1, N_2, \ldots, N_M)$ with N_i indicating the number of SSs currently using class i traffic in a cell. Then, the state space for each cell Ω consists of all allowable states such that $\mathbf{N} \cdot \mathbf{B} \leq C$. Define a vector $\mathbf{S} = (S_1, S_2, \ldots, S_M)$ where $1 \leq S_i \leq N_i$ and S_i represents the number of SSs located in A_i ($i \geq 2$) area (in this study, we assume an SS is not allowed to have multiple sessions). So, the SSs in the vector can be redirected to other BSs for load balancing if one of the BSs has enough bandwidth to accommodate it. That is, if an SS belongs to S_i the CS may request it to change BSs from the current cell to an adjacent cell so that the amount of B_i can be secured in the current cell. The CS must maintain information described above for each cell as well as lists of SS ID's in S_i as indicated in Table I. Note that $\mathbf{N} \cdot \mathbf{B}$ indicates the bandwidth being used by SSs in a cell while $\mathbf{S} \cdot \mathbf{B}$ represents that employed by SSs in the overlapping region between the cell and its neighboring cells. Thus, the amount of bandwidth that can be redirected to other cells is always less than or equal to $\mathbf{S} \cdot \mathbf{B}$.

C. Admission Control

After receiving a bandwidth request (B_{req}) from a new SS the CS carries out the following admission control procedure. It consists of three steps:

- **A1)** If $\mathbf{N} \cdot \mathbf{B} \leq C - B_{req}$, allocate bandwidth B_{req} to the SS and the procedure ends. Otherwise go to **A2)**.
- **A2)** If the SS is located in overlapping area, do the same test **A1)** for other accessible cells one by one. The procedure stops if one of them is found to have enough bandwidth to accommodate B_{req}. In this case the bandwidth is allocated to the SS and the procedure ends. If there is no such cells, then go to **A3)**.
- **A3)** If $\mathbf{S} \cdot \mathbf{B} < B_{req}$, the request is blocked; otherwise, the CS makes the required bandwidth available for the SS by requesting some SSs belonging to $\mathbf{S}$ to change BSs.

When only step **A1)** is adopted for admission control it corresponds to so-called fixed allocation **(FA)** scheme. If steps **A1)** and **A2)** are utilized it is called directed retry **(DR)** scheme in this paper as it is similar to that proposed in [6], [7] for mobile cellular networks. **DR** allows a new call that cannot be served at one BS to attempt access via a nearby alternate BS. Step **A3)** contains an interesting issue of choosing SSs in $\mathbf{S}$ such that the required bandwidth is made available in the cell. Since the state space of $\mathbf{S}$ contains $\prod_{j=1}^{M}(S_j + 1)$ number of possible vectors, we need a criteria when choosing one of them to secure the required bandwidth. We take two factors into account in this step. First, bandwidth obtained by sending SSs to other cells should be as close to B_{req} as possible because impressing large bandwidth on neighboring cells might cause higher connection blocking in the future. Second, the number of SSs involved in the procedure should be taken into account since the more SSs the larger overhead is introduced to change their connection to other cells. To this end, a cost function is introduced as follows.

Suppose one SS in $\mathbf{S}$ having bandwidth B_k greater than or equal to B_{req} is chosen. If there is a neighboring cell with enough bandwidth for the SS it becomes an *upward candidate* for B_{req}. In this case the cost, called *upward cost*, is defined as

$$B_k - B_{req} \tag{1}$$

When the SS is redirected to other cell an unnecessary traffic load of $(B_k - B_{req})$ is impressed on the neighboring cell that might increase traffic blocking in the cell. On the other hand, when we choose SSs in $\mathbf{S}$, each having a bandwidth lower than B_{req}, multiple SSs that could be taken over to neighboring cells must be selected so that their total aggregate bandwidth is greater than or equal to B_{req}. They constitute a vector ($\mathbf{S}^*$) such that $\mathbf{S}^* \cdot \mathbf{B} \geq B_{req}$. In this case the cost is defined as

$$(|\mathbf{S}^*| - 1) + (\mathbf{S}^* \cdot \mathbf{B} - B_{req}) \tag{2}$$

where $|\mathbf{S}^*|$ is the number of SSs in $\mathbf{S}^*$. The SSs are called *downward candidates* and the cost is referred to as *downward cost*. The first term takes the number of SSs involved into consideration since the overhead for commanding them to change BSs increases with it. The second term explains the difference between the resultant bandwidth and the required bandwidth. It should be emphasized that an SS in overlapping area can be either an upward or a downward candidate only when one of the accessible cells other than the cell serving it has enough bandwidth to accommodate it. Now the step **A3)** is modified as follows:

- **A3')** If $\mathbf{S} \cdot \mathbf{B} < B_{req}$ the request is blocked; otherwise, the CS requests some SSs belonging to $\mathbf{S}$ to change BSs making bandwidth available for the SS such that the chosen SS(s) minimizes the cost defined above.

TABLE I

TRAFFIC CLASS, BANDWIDTH, AND SOME VECTORS FOR A CELL THAT THE CS MAINTAINS FOR LOAD BALANCING

Traffic class	1	2	$\cdots$	k	$\cdots$	M
Bandwidth of class ($\mathbf{B}$)	B_1	B_2	$\cdots$	B_k	$\cdots$	B_M
$\mathbf{N}$	N_1	N_2	$\cdots$	N_k	$\cdots$	N_M
$\mathbf{S}$	S_1	S_2	$\cdots$	S_k	$\cdots$	S_M
SS ID in $\mathbf{S}$	$SS_{1,1}$ $\cdots$ SS_{1,S_1}	$SS_{2,1}$ $\cdots$ SS_{2,S_2}	$\cdots$	$SS_{k,1}$ $\cdots$ SS_{k,S_k}	$\cdots$	$SS_{M,1}$ $\cdots$ SS_{M,S_M}

D. Algorithm for Choosing Candidates

We propose a simple algorithm for choosing candidate SSs in admission control step **A3')** assuming the required bandwidth B_{req} is in class i (i.e., $B_{req} = B_i$).

C1) If S_i is not empty, examine it in search of an SS that can be taken over to a neighboring cell. If there is such an SS, terminate the algorithm (in this case the cost is zero). Otherwise, go to step **C2)**.

C2) In this step an upward candidate is chosen. The CS scans S_k from $k = i + 1$ to $k = M$. If S_k is not empty, search an upward candidate in S_k. Once an upward candidate is found, the search stops and the upward cost becomes $B_k - B_i$. When there is no such candidate the upward cost is set to ∞ and go to step **C3)**.

C3) Downward candidates are selected in this step.

1) The CS scans S_k from $k = i - 1$ to $k = 1$ and when S_k is not empty it puts the SSs in S_k that can be taken over to other cells into the *candidate list*. This operation stops either when total aggregate bandwidth of the SSs in the list is greater than or equal to B_i or when there are no such SSs in all $S_k, k = i-1, \ldots, 1$. In the latter case the downward cost is set to ∞ and go to step **C4)**.
2) When the above procedure stops at $k = i - l, l = 1, \ldots, i - 2$, the CS investigates S_k from $k = i - l$ to $k = 1$. In the process, if S_k is not empty one SS in S_k that can be a downward candidate is replaced with one of the SSs in the *candidate list* with maximum possible bandwidth such that total aggregate bandwidth of the SSs in the list is greater than or equal to B_i. This procedure continues until $k = 1$ and the downward cost is calculated. The SS in the list constitutes $\mathbf{S}^*$ in Eq. (2). Note that by replacing an SS in the list with an SS with lower bandwidth the downward cost is reduced. Go to step **C4)**.

C4) Compare upward cost and downward cost and select the least. If both costs are ∞, the connection request is blocked.

Unless the cost chosen in step **C4)** is not ∞, the corresponding candidate(s) will be requested by the CS to change BSs in order to secure bandwidth for the requesting SS. When there is no such candidates the request is blocked.

TABLE II

AN EXAMPLE OF TABLE I

Traffic class	1	2	3	4	5	6	7	8	9	10
Bandwidth of class ($\mathbf{B}$)	1	2	3	4	5	6	7	8	9	10
$\mathbf{N}$	3	5	1	2	0	0	2	0	0	2
$\mathbf{S}$	2	1	1	0	0	0	0	0	0	1

Table II shows an example of traffic class, its associated bandwidth ($\mathbf{B}$) and vectors $\mathbf{N}$ and $\mathbf{S}$. We assume here that C is 60, neighboring cells have enough bandwidth to accommodate SSs in $\mathbf{S}$ and $B_{req} = B_5$ from a SS in A_1. As $\mathbf{N} \cdot \mathbf{B} > C - B_5$ and the SS is located in A_1 the first two step **A1)** and **A2)** fail. However, since $\mathbf{S} \cdot \mathbf{B} > B_5$ we see there is a solution. Following the procedure described in **C1)** and **C2)** the upward candidate belonging to S_{10} is found, so the upward cost is five. On the other hand, the downward candidate $\mathbf{S}^*$ is obtained using **C3)** resulting in $\mathbf{S}^* = (0, 1, 1, 0, \ldots, 0)$ and the downward cost is one. Therefore, downward candidate is chosen and the CS requests the SSs in $\mathbf{S}^*$ to change BSs. After that the requesting SS is allocated bandwidth B_5.

III. PERFORMANCE EVALUATION

In this section a simulation study for the proposed system is described. Performance of the proposed load balancing **(LB)** algorithm is compared with **FA** and **DR** schemes described in admission control subsection.

A. Simulation Assumptions and Parameters

For a simulation study we consider 16 cell model with wraparound to avoid edge effects as shown in Fig. 2. SSs are generated according to traffic model and disappear after finishing lifetime. TDMA system is assumed where one frame is composed of 80 slots and each slot fits into one packet. Four different CBR traffic classes are considered, each requiring one, four, eight, and ten slots per frame, respectively. Traffic generation follows Poisson process and lifetime is exponentially distributed with a mean of 5000 frame time.

A frequently used measure is offered load per cell (L), which is defined as connection generation rate $\times$ connections' bandwidth $\times$ average connection lifetime, i.e.,

$$L = \lambda \times (1 \cdot P_1 + 4 \cdot P_2 + 8 \cdot P_3 + 10 \cdot P_4) \times 5000 \cdot t_f \quad (3)$$

where λ is connection generation rate and P_i is the probability of generating class i traffic with $\sum_i P_i = 1$ and t_f is the frame time. The physical meaning of the offered load per cell is the total bandwidth required on average to support all existing connections in a cell. The percentages of each class, $P_i, i = 1, 2, 3, 4$ in Eq. (3), are 67.8 %, 16.9 %, 8.5 % and 6.8 %, respectively, such that the traffic load of each class is the same.

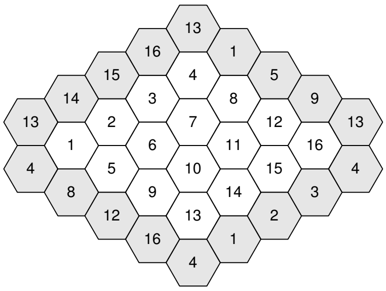

Fig. 2. 16 cell model with wraparound for simulation.

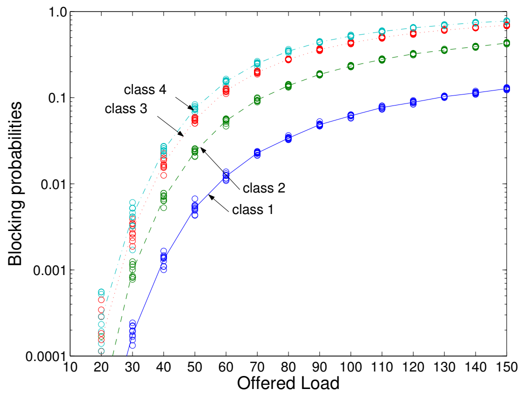

Fig. 3. Blocking probabilities of fixed assignment **(FA)** scheme when $R_r = 1.3$.

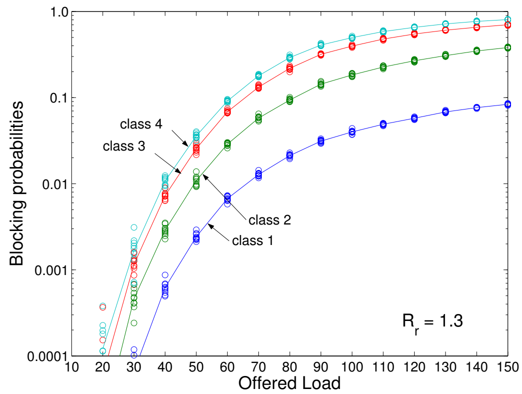

Fig. 4. Blocking probabilities of directed retry **(DR)** scheme when $R_r = 1.3$.

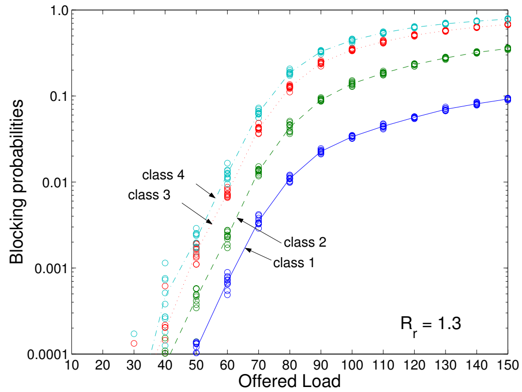

Fig. 5. Blocking probabilities of the proposed load balancing **(LB)** scheme when $R_r = 1.3$.

B. Numerical Results

Blocking probabilities for **FA**, **DR**, and **LB** are shown in Fig. 3, 4, and 5, respectively, when $R_r = 1.3$. Note that when $R_r = 1.3$ the proportions of A_1, A_2 and A_3 are about 25.9 %, 43.8 % and 30.3 %, respectively; that is, more than 74.1 % of SSs are located in overlapping area on average. Each simulation was run for at least 8×10^6 slots with 10 % of warming-up period. After 10 times of simulation run average value was taken. With low and medium offered load **LB** scheme performs much better than the other two schemes **FA** and **DR** in terms of connection blocking probability. As traffic grows heavy the performance difference between the three schemes becomes close due to shortage of bandwidth for load balancing. We also see that SSs with higher bandwidth requirement have higher blocking probabilities than those of lower bandwidth requirement as can be expected. To be fair we need to introduce a threshold for each class such that SSs with varied bandwidth requirements experience the similar blocking probabilities. In addition, as indicated in [8], SSs in $A_n, n \geq 2$ region are more likely to find capacity than those in A_1. Thus, another sort of threshold to SSs in overlapping region is necessary for the system to be fair. However, since the distance from BS to SSs in overlapping region is greater than that for SSs in A_1 SSs in overlapping region have higher bit error rate or packet error rate when transmit power is the same as indicated in [9]. Concerning this issue further study is under way.

Fig. 6 shows channel utilization of three schemes when $R_r = 1.3$, representing **LB** outperforms other two ones. Cost defined in Eq. (1) and (2) associated with **LB** is depicted in Fig. 7 when $R_r = 1.3$. It grows until offered load is 90, and after that it decreases with offered load. This can be explained that as traffic load increases it becomes harder to find candidate SS's that could be redirected to neighboring cells due to lack of resources.

Blocking probabilities of **LB** when $R_r = 1.0$ and $R_r = 1.5$ are shown in Fig. 8 and 9, respectively. Since when $R_r = 1.0$ the percentage of overlapping area is only 20.9 % small improvement is observed compared to **FA** shown in Fig. 3. With $R_r = 1.5$ although most of the area is overlapped (92.7 %) small improvement is seen compared to the case of $R_r = 1.3$ depicted in Fig. 5.

Table III shows connection blocking probabilities when nonuniform traffic load is considered; in particular, the offered load of cell 10 is 100 and those of other cells are 50. **LB** indicates higher performance compared to the other schemes not only in cell 10 but also in the cells located in the first tier of cell 10.

IV. Conclusions and Future Work

Centralized wireless network architecture requires us to reconsider most of system resource management issues since

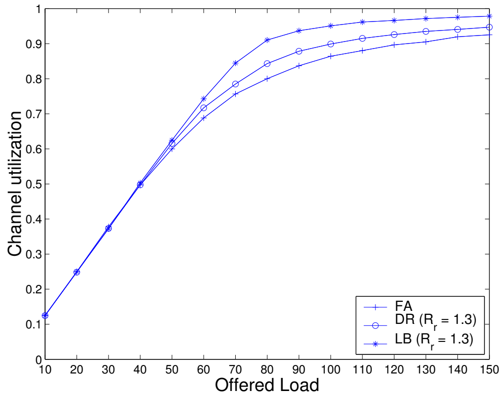

Fig. 6. Channel utilization when R_r is 1.3.

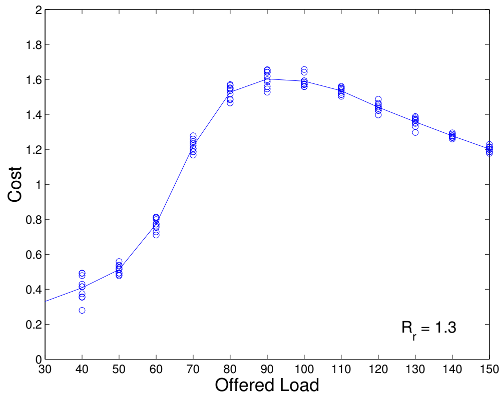

Fig. 7. Cost vs. offered load of **LB**.

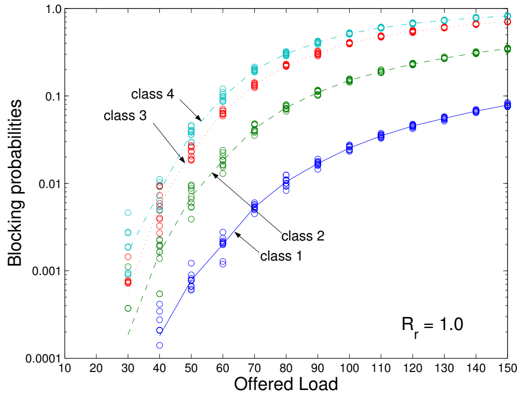

Fig. 8. Blocking probabilities of **LB** when $R_r = 1.0$.

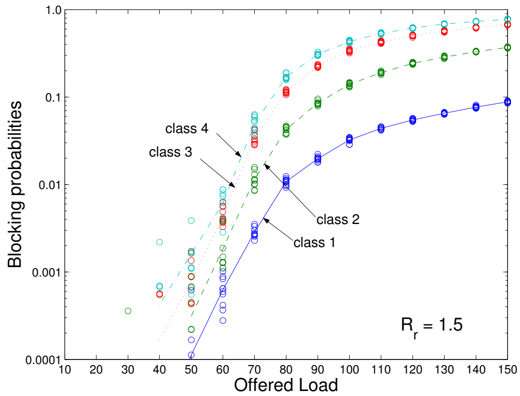

Fig. 9. Blocking probabilities of **LB** when $R_r = 1.5$.

research efforts for them has mainly been proposed for distributed wireless architecture. Dynamic channel assignment, dynamic transmit power control, load balancing and mobility management are representative of such management issues. In this paper we have proposed a dynamic load balancing algorithm exploiting centralized control capability for multi-class fixed broadband wireless access networks. The algorithm has been compared to conventional ones through a simulation study, indicating its capability of efficient system resource. Future works will involve a dynamic load balancing scheme that takes into account interference between BSs, distance between user and BSs and user mobility.

TABLE III

BLOCKING PROBABILITIES WHEN OFFERED LOAD IS 100 FOR CELL 10 AND 50 FOR OTHER CELLS WITH $R_r = 1.3$.

Cell	Class	FA	DR	LB
10	1	0.063741	0.022373	0.000141
	2	0.236932	0.088398	0.001521
	3	0.431905	0.166851	0.006396
	4	0.516552	0.209347	0.013251
Cells in the first tier	1	0.005392	0.003074	0.000300
	2	0.024097	0.015454	0.000719
	3	0.054493	0.037988	0.002842
	4	0.073799	0.050411	0.004095

REFERENCES

[1] H. Al-Raweshidy and S. Komaki, editors, *Radio over Fiber Technologies for Mobile Communications Networks*, Norwood: Artech House, 2002.

[2] A. Hills and B. Friday, “Radio Resource Management in Wireless LANs,” *IEEE Commun. Mag.*, pp. S9–S14, Dec. 2004.

[3] H. B. Kim, H. Woesner, and A. Wolisz, “A Medium Access Control Protocol for Radio over Fiber Wireless LAN operating in the 60-GHz band,” in *Proc. 5th European Personal Mobile Commun. Conf.*, pp. 204–208, Apr. 2003.

[4] H. B. Kim, M. Emmelmann, B. Rathke, and A. Wolisz, “A Radio over Fiber Network Architecture for Road Vehicle Communication Systems,” in *Proc. IEEE VTC 2005 Spring.*, May. 2005, Stockholm, Sweden.

[5] N. Srivastava and S. S. Rappaport, “Models for overlapping coverage areas in cellular and micro-cellular communication systems,” in *Proc. IEEE Globecom 1991*, vol. 2, pp. 890–894, Dec.1991.

[6] B. Eklundh, “Channel Utilization and Blocking Probability in a Cellular Mobile Telephone System with Directed Retry,” *IEEE Trans. Commun.*, vol. 34, no. 4, pp. 329–337, Apr. 1986.

[7] J. Karlsson and B. Eklundh, “A Cellular Mobile Telephone System with Load Sharing- An Enhancement of Directed Retry,” *IEEE Trans. Commun.*, vol. 37, no. 5, pp. 530–535, May. 1989.

[8] X. Lagrange and B. Jabbari, “Fairness in Wireless Microcellular Networks,” *IEEE Trans. Veh. Technol.*, vol. 47, no. 2, pp. 472–479, May. 1998.

[9] G. Bianchi and H. Tinnirello, “Improving Load Balancing mechanisms in Wireless Packet Networks,” *Proc. IEEE ICC 2002*, pp. 891–895, 2002.

[10] T. Chu and S. S. Rappaport, “Overlapping Coverage with Reuse Partitioning in Cellular Communication Systems,” *IEEE Trans. Veh. Technol.*, vol. 46, no. 1, pp. 41–54, Feb. 1997.