Integrated Cellular and Ad Hoc Relay (iCAR) Systems: Pushing the Performance Limits of Conventional Wireless Networks*

Swades De
EE Department
State University of New York at Buffalo
Buffalo, NY 14260, USA
swadesd@cse.buffalo.edu

Ozan Tonguz
ECE Department
Carnegie Mellon University
Pittsburgh, PA 15213, USA
tonguz@ece.cmu.edu

Hongyi Wu   Chunming Qiao
CSE Department
State University of New York at Buffalo
Buffalo, NY 14260, USA
{*hongyiwu, qiao*}*@cse.buffalo.edu*

Abstract

Performance of iCAR system - a new load balancing scheme in wireless networks - is analyzed. Traffic capacity enhancement of the iCAR system with respect to the conventional cellular system, without any load balancing, is evaluated. It is shown that, with a moderate amount of relay coverage, perfect load balancing can be achieved, thus enabling the system to support maximum possible traffic intensity for a given grade of service.

1 Introduction

Traditional cellular systems have been very successful in providing voice services since the first analog system was introduced about fifteen years ago. In the last decade, with the unprecedented increase in demand for personal mobility and dependence on personal communications, both the number of subscribers and the amount of wireless traffic have surged at an exploding speed. With the advent of Internet, especially the wireless access to the Internet, wireless data traffic is expected to exacerbate the demand for bandwidth. The carriers and infrastructure providers now face a major challenge in meeting the increased bandwidth demand of mobile Internet users.

At the same time, various efforts in providing different access services such as wireless LANs, ad hoc networks, Bluetooth and home RF networks, are further stimulating the growth of wireless traffic and the requirement for a ubiquitous wireless infrastructure. Moreover, continued proliferation of these services will call for interoperability between heterogeneous networks such as ad hoc and cellular systems. In addition, such an interoperability will create heavier traffic in cellular systems as more and more traffic from wireless LANs, ad hoc networks and Bluetooth devices, will be carried by the cellular infrastructure.

*This research is in part supported by NSF under the contract ANIR-ITR 0082916.

For reasons cited above and the fact that the traffic in future cellular systems will be more bursty and unevenly distributed than conventional voice traffic, it is anticipated that *congestion* will occur in peak usage hours even in the next generation (e.g., 3rd generation or 3G) systems, which will have increased capacity. By congestion, we mean that in some cells, data channels (DCHs) are less frequently available, thereby deteriorating the grade of service (GoS) in those cells to a level below a prescribed threshold (e.g., the GoS above 2%). Note, however, that control channels (CCHs) for signaling (or paging) may still be accessible by all mobile hosts (MHs) in a congested cell.

Presence of *unbalanced traffic* will exacerbate the problem of limited capacity in existing wireless systems. Specifically, some cells may be heavily congested (called *hot spots*), while the other cells may still have enough available DCHs. In other words, even though the traffic load does not reach the maximum capacity of the entire system, a significant number of calls may be blocked and dropped due to localized congestion. Since the locations of hot spots vary from time to time (e.g., downtown areas on Monday morning, or amusement parks in Sunday afternoon), it is difficult, if not impossible, to provide the guarantee of sufficient resources in each cell in a cost-effective way. Congestion due to unbalanced traffic can be a real problem in wireless networks. For example, when providing emergency telecommunications at a disaster site, due to heavy cellular traffic demand, severe congestion may be experienced by critical disaster relief officials when communications are needed the most [1]. Increasing bandwidth of a cellular system (e.g., the number of DCHs in each cell) can increase the system capacity but not the efficiency in dealing with the time-varying unbalanced traffic.

Recently, a novel approach has been proposed in [3],[4], which shows a direction of how to evolve from the existing, heavily-invested cellular infrastructure to next generation wireless systems that scale well with the number of mobile hosts and, in particular, overcome the congestion by *dynamically balancing the load* among different cells in a cost-effective way. The scheme combines conventional cellular technology and ad hoc wireless networking technology. The basic idea of the proposed system, called iCAR (integrated Cellular and Ad hoc Relay), is to place a number of ad hoc relay stations (ARSs) at strategic locations, which can be used to relay signals between MHs and base stations. By using ARSs, it is possible to divert traffic from one (possibly congested) cell to another (non-congested) cell. This helps to circumvent congestion, and makes it possible to maintain (or hand-off) calls involving MHs that are moving into a congested cell, or to accept new call requests involving MHs that are in a congested cell. Although this paper focuses only on the issues related to load balancing, there are many other benefits of the proposed iCAR system. For example, the ARSs can, in a flexible manner, extend the cellular system's coverage (similar to the wireless routers used in the Rooftop system[5]), and provide interoperability between heterogeneous systems (by connecting ad hoc networks and wireless LANs to Internet for example). Additional benefits include enhanced reliability (or fault-tolerance) of the system, and potential improvement in MHs' battery life and transmission rate.

In the multihop cellular systems approach presented in [2], relaying is performed by MHs, and thus that approach shares many disadvantages in terms of security (authentication, privacy), billing, and mobility management (of the MHs) with mobile ad hoc networks as discussed in [4]. In addition, the main goal of the multihop cellular systems is to reduce the number of base transceiver systems (BTSs) or the transmission power of each BTS, but it can no longer guarantee a full coverage of the area. In fact, even in the ideal case where every MH (in an area not covered by any BTS) can find a relaying route (through other MHs), the multihop approach will neither increase the system capacity nor decrease the call blocking/dropping probability, unless a large percentage of the calls are intra-cell calls (i.e., calls whose source and destination are in the same cell), which usually is not the case in practice.

Note that the proposed relaying through ARSs is useful in any cellular system where congestion may occur, even though a call may not be allocated a dedicated DCH all the time (or in other words, during the entire call duration). Also, if one simply treats the ISM band as an additional set of channels that can be used in a cellular system (by, e.g., modifying each BTS so that it is equipped with the R-interface as well), one will not be able to balance loads among cells or to eliminate congestion in hot-spot cells via relaying. Other approaches such as those using cell splitting and sectorization can not serve as substitutes either, although they may be used in conjunction with the approach proposed in [4].

Our objective here is to analyze load balancing performance of the recently proposed iCAR system. We quantify the steady-state performance of iCAR system under *idealized wireless channel conditions*, where the fading effects are assumed non-existent and the usage of ISM band poses no restriction. We show that, in the three-tier example network studied in this paper, the iCAR system increases channel capacity of a congested cell by approximately 70%.

Remainder of the paper is organized as follows. Section 2 reviews the principle of operation and main benefits of the proposed iCAR system. In Section 3, we present the system performance analysis of iCAR. Section 4 contains the numerical results. Section 5 concludes the paper.

2 An Overview of iCAR System

In this section, we describe briefly the principle of operation and the main benefits of iCAR system (see [4] for more details). To simplify the following presentation, we will focus on cellular systems where each BTS is controlled by a Mobile Switching Center (MSC) [6],[7].

Each ARS has two air interfaces, the C (for cellular) interface for communications with a BTS and the R (for relaying) interface for communicating with an MH or another ARS. Also, MHs should have two air interfaces: the C interface for communicating with a BTS, and the R interface for communicating with an ARS. In addition, each ARS is under the control of a MSC, and has limited mobility. Such a feature is important to ensure that a relaying route can be set up fast and maintained with a high degree of stability. Routing in the proposed system is similar to that of having a hybrid (both hierarchical and flat) structure in [8] for efficient routing and hand-offs in mobile ATM networks. The difference between the two is that in the latter, path extension (or relay) is between two (fixed) BTSs through direct wired links. The R interface (as well as the medium access control (MAC) protocol used) is similar to that used in wireless LANs or ad hoc networks (see for example [9]-[18]). Note that because multiple ARSs can be used for relaying, the transmission range of each ARS using its R interface can be much shorter than that of a BTS, which implies that an ARS can be much smaller and less costly than a BTS. At the same time, it is possible for ARSs to communicate with each other and with BTSs at a higher data rate than MHs can, due to limited mobility of ARSs and specialized hardware (and power source).

In the iCAR system relaying occurs even without occurance of congestion in the network, such that whenever there is a difference in traffic pattern among neighboring cells the relays are activated to mitigate the difference. Interference in cellular band due to channel borrowing is avoided in this scheme. However, one still needs to take care of the interference issues in the ISM band.

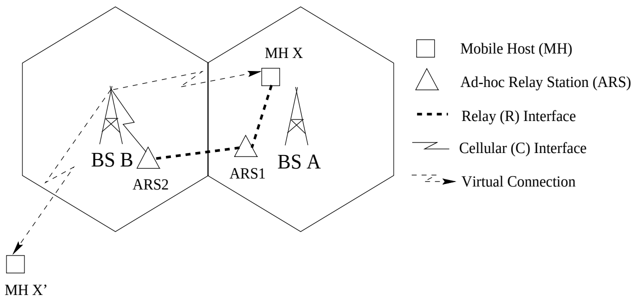

Figure 1: Primary relaying. Mobile host X (*MH X*) in cell A operates on a channel from base station B via ISM band relays.

Primary Relaying: In an existing cellular system, without any load balancing strategy incorporated, if MH X in Fig. 1 is involved in a new call (as a caller or callee) but it finds no DCH in cell A at that moment, the new call will be blocked. In the iCAR system, MH X in cell A, can *switch-over* to the R interface to communicate with an ARS in a neighboring cell (cell B, which is less congested), possibly through other ARSs in cell A (see Fig. 1 for an example), and thus the call can be served directly by relaying. We call this strategy *primary relaying*. The process of changing over from C interface to R interface (or vice versa) is referred as switching-over, which is similar to (but different from) frequency-hopping [6],[19],[20]. A relaying route between MH X and its corresponding (i.e., caller or callee) MH X' can also be established, in which case, both MHs need to switch-over from their C interfaces to their R interfaces, even though the probability of this event is typically very low.

Secondary Relaying: If primary relaying is not possible, because, for example in Fig. 1, ARS 1 is not close enough to MH X to be a proxy (and there are no other nearby ARSs), then one may resort to *secondary relaying* so as to free up a DCH from BTS A for MH X. An example is given in Fig. 2(a), where MH Y denotes any MH in cell A which is currently involved in a call. One may establish a relaying route between MH Y and BTS B (or any other neighboring less congested cell). In this way, after MH Y switches-over, the DCH freed by MH Y can now be used by MH X. Note that, since the probability of finding an on-going call covered by an ARS is much higher than that of a blocked call, the likelihood of secondary relaying is much higher than that of primary relaying. In addition, although the concept of having an MH-to-MH call via ARSs only (i.e., no BTSs are involved) is similar to that in ad hoc networking, a distinct feature (and advantage) of the proposed integrated system is that an MSC can perform (or at least assist in performing) critical call management functions such as authentication, billing, and locating the two MHs and finding and/or establishing a relaying route between them, as mentioned earlier. Such a feature is also important to ensure that switching-over of the two MHs (this concept is not applicable to ad hoc networks) is completed fast enough so as not to disconnect the on-going call involving the two MHs or not to cause severe quality of service (QoS) degradation (even though the two MHs may experience a "glitch" or jitter).

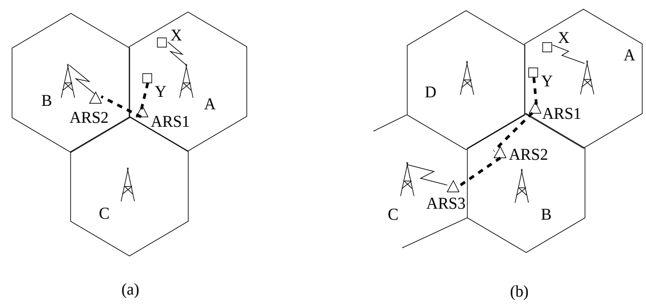

Figure 2: (a) Secondary relaying. *MH X* in cell A operates on a freed channel from *MH Y*. *MH Y* now operates on a borrowed channel from cell B via ARS_1 and ARS_2. (b) Cascaded-secondary relaying. *MH X* operates on a freed channel from *MH Y* in cell A. *MH Y* now operates on a borrowed channel from cell C via relays ARS_1, ARS_2, and ARS_3.

Cascaded-Secondary Relaying: If neither primary relaying (as shown in Fig. 1), nor basic secondary relaying (as shown in Fig. 2(a)) works, the new call may still be supported. As shown in Fig. 2(b), one may apply the basic secondary relaying strategy twice in cascade, to relay an on-going call from a host (MH Y) to a cell (cell C in Fig. 2(b)). In this way, the freed channel can be allocated to a new host (MH X), while a borrowed channel from cell C will be allocated to the MH Y. We call this strategy *cascaded relaying*. Note that in this case cell B in our example (see Fig. 2) does not lose any channel capacity.

In addition to the above relaying strategies, one design issue that is critical in iCAR is the number and placement of ARSs[4].

3 Performance Analysis

Given the above description of the iCAR system, in this section we analyze its performance. The analysis shows how the variation of ARS coverage (i.e., p) affects the system load balancing performance. It also shows the maximum possible capacity gain that can be achieved in the proposed iCAR system. Although in reality the traffic distribution in different cells can be uneven, for analytical tractability we consider a 3-tier cellular structure, as shown in Fig. 3, where the most congested cell (A) is surrounded by less-congested tier B cells, which, in turn, are surrounded by even lesser-congested tier C cells. The analysis mainly takes into account the traffic imbalance in different cells and assumes smooth (Poissonian) traffic arrival and independent (exponentially distributed) service process, i.e., Erlang-B (*M/M/m/m* queuing) traffic model [6]. Also, the ISM band channel capacity for relaying traffic is assumed to be sufficient so that whenever a mobile host is reachable to an ARS, relaying is possible.

We provide here a steady-state solution for the traffic intensities achieved after load balancing through relaying, which combines the effects of primary and secondary relaying. The strategy for load balancing in the proposed iCAR (and hence the analytical framework outlined below) is as follows: Primary relaying operates on new calls in a cell to reduce the traffic intensity in that cell to exceed the average traffic intensity of the network. However, due to limited ARS coverage, its effect is limited, because it can only operate on a fraction p (which is the normalized ARS coverage area in a cell) of blocked calls. Secondary relaying, on the other hand, operates on ongoing calls within the ARS coverage area in that cell so that the remaining "heat" (excess traffic) can be distributed among the neighboring cells. One has to devise this strategy; otherwise, if one tries to relay every blocked call from a "hot" cell, that cell may become "cooler" than the average "temperature" of the network, and the surrounding cells, which take those extra calls, are going to be "hotter" than the average "temperature" of the network. In other words, this strategy is required for proper load balancing.

We focus on a three-tier system where that the "hot" cell (cell A) is surrounded by "cooler" (in the sense of traffic intensity) cells (i.e., tier B and tier C cells), as depicted in Figure 3 (because the effect of relaying become less significant beyond tier C, and similar techniques can be applied to a bigger system as well). We also assume that spatial distribution of calls in a cell is uniform, so that a call is covered by an ARS with probability p. Tier B and tier C cells have stable traffic patterns, T_b and T_c (in *Erlang*), respectively. Consider the scenario wherein cell A traffic (T_a *Erlang*) is growing, thus causing a growing impact on surrounding tier B and tier C cells. A relatively "hot" cell spreads out its traffic to its surrounding relatively "cold" cells through primary and secondary relaying, respectively. Traffic is not spread to an equally (or more) loaded surrounding cell. We do not consider the cascaded-secondary relaying scheme in our analysis, as in our three tier cell model it gives little improvement over secondary relaying.

Since cellular band interference due to channel borrowing is absent in the iCAR system, given T_a, T_b, and T_c, if perfect load balancing is achieved, steady-state traffic intensity per cell will be given by

$$T_f = \frac{T_a + 6T_b + 12T_c}{19}. \tag{1}$$

The excess traffic in a cell, $T_i - T_f$ (where $i = a,\ b,\ c$), will determine the amount of load balancing that will be

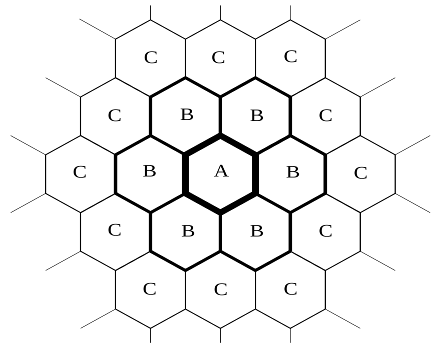

Figure 3: The cell model considered in our analysis. The cell with the thickest boundary indicates the most congested cell, whereas the cell with the thinnest boundary indicates the least congested one.

required in that cell. In our model, $T_a > T_f$, $T_c < T_f$, and T_b could be larger, equal, or smaller than T_f.

Excess traffic to be relayed from cell A to neighboring tier B cells is $(T_a - T_f)$. Traffic that can be relayed in cell A is pT_a, out of which maximum traffic that can be served through primary and secondary relaying is $pT_a(1 - B_b^r)$, where B_b^r is the steady-state blocking probability in tier B cells. For certain traffic patterns, beyond a specific p value, relaying more than ΔT_a (where $0 < \Delta T_a \leq pT_a$) amount of traffic from cell A may cause cell A to become "cooler" than the surrounding tier B cells, which could lead to *reverse cooling effect*[1]. This introduces new traffic imbalance, which is not acceptable. Therefore, maximum allowable traffic out of $pT_a(1 - B_b^r)$ is $\Delta T_a(1 - B_b^r)$. So, total traffic served through relaying in cell A is

$$R_a = \min\left\{(T_a - T_f),\ \Delta T_a(1 - B_b^r)\right\}. \tag{2}$$

Therefore, in the steady-state, remaining traffic in cell A after relaying is:

$$T_a^r = T_a - R_a. \tag{3}$$

Hence, steady-state call blocking probability in cell A due to relaying is obtained as

$$B_a^r = \frac{\left(T_a^r\right)^m / m!}{\sum_{i=0}^{m} \left(T_a^r\right)^i / i!} \triangleq f(T_a^r,\ B_b^r,\ m), \tag{4}$$

where m is the number of cellular band channels per cell.

Since the traffic in a cell is uniformly distributed and the call coverage by an ARS in a cell is uniform, without loss of generality, we assume that relayed calls from cell A will load all six tier B cells equally. Thus, all tier B cells being

[1]This situation may arise when the initial traffic loads in tier C cells are very low, whereas traffic loads in cell A, and tier B cells are almost the same and at a much higher level, and the fractional ARS coverage, p, has an intermediate value; i.e., p is such that only partial load balancing is achievable.

Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS-35'02)
0-7695-1435-9/02 $17.00 © 2002 IEEE

equally loaded, they will not exchange traffic through relaying. So, a tier B cell can only relay traffic to its surrounding tier C cells (since these cells have less call blocking probability). Traffic relayed from cell A to each of surrounding tier B cells is $\frac{R_a}{6}$. So, total traffic in cell B, including relayed traffic from cell A, is:

$$T_b^{'} = T_b + \frac{R_a}{6}. \tag{5}$$

This, however, will be reduced due to relaying to the surrounding ("cooler") tier C cells. Excess traffic in a cell B, that has to be relayed for perfect load balance is $(T_b^{'} - T_f)$. Each tier B cell is surrounded by three tier C cells. Since spatial distribution of calls in a cell is uniform, only 50% of traffic covered by ARSs in a cell B can be relayed to tier C cells. Therefore, maximum traffic that can be served through primary and secondary relaying is $\frac{p}{2}T_b(1 - B_c^r)$, where B_c^r is the steady-state blocking probability in tier C cells. Again, as in the case of relaying from cell A to tier B cells, it may happen that, for certain traffic patterns, beyond ΔT_b (where $0 < \Delta T_b \leq \frac{p}{2}T_b$) amount of traffic from tier B cells, there may be undesired reverse cooling effect[2]. Taking this factor in account, total traffic served through relaying in a tier B cell is:

$$R_b = \min\left\{(T_b^{'} - T_f),\ \Delta T_b(1 - B_c^r)\right\}. \tag{6}$$

Therefore, in the steady-state, remaining traffic in a tier B cell after relaying is:

$$T_b^r = T_b^{'} - R_b, \tag{7}$$

from which one can obtain steady-state call blocking probability in a tier B cell due to relaying as $B_b^r = f(T_b^r,\ B_c^r\ m)$.

Again, assuming uniform traffic distribution, a tier B cell will relay equal amount of traffic through each of the three boundaries connected to the tier C cells. Thus, traffic relayed through each of the three boundaries from a tier B cell is $\frac{R_b}{3}$. Traffic received from tier B cells by alternate cells in tier C are different. A tier C cell, which has only one boundary to the tier B cells, will receive $\frac{R_b}{3}$ amount of traffic, whereas a tier C cell which has two boundaries to tier B cells will receive $\frac{2R_b}{3}$ amount of traffic. From Fig. 3, observe that the alternate C cells will have such an imbalance in the amount of received extra traffic. Therefore, although originally each of tier C cells were carrying equal amount of traffic, due to the relayed traffic, there will be imbalance between neighboring tier C cells, and hence exchange of traffic. In the steady-state, excess amount of traffic that each of the tier C cells will have, is $\frac{1}{2}\left(\frac{R_b}{3} + \frac{2R_b}{3}\right) = \frac{R_b}{2}$. So, in the steady-state, total traffic carried by each of the tier C cells is

$$T_c^r = T_c + \frac{R_b}{2}, \tag{8}$$

which makes the steady-state call blocking probability in a tier C cell as $B_c^r = f(T_c^r,\ B_c^r,\ m)$.

Note that the Eqs. (2) and (6), and hence steady-state call blocking probabilities in three tiers are inter-related. However, they remain unequal until the ARS coverage (p) is not sufficient enough to perfectly balance the traffic load in different tiers.

By numerically solving Eqs. (2) and (6), steady-state call blocking probabilities in different tiers and the traffic capacity gain in cell A are obtained. One way to verify the analytical results of steady-state iCAR system performance is the iterative approach. We compute the call blocking probabilities iteratively, by relaying δT_a and δT_b amount of traffic from cell A and a tier B cell, respectively, and making δT_a and δT_b reasonably small. The algorithm for the iterative approach is presented in the next section. The numerical solution and the iterative approach yield identical results.

4 Numerical Results

In this section, we quantify the iCAR system performance. We are primarily interested in obtaining the system capacity enhancement through the proposed dynamic load balancing scheme. The iterative computation algorithm is given below, where the respective instantaneous call blocking probabilities in cells A, B, and C are denoted as B_a, B_b, and B_c.

1. Initialize $R_{a,max} = 0$, $R_{b,max} = 0$
2. Read T_a, T_b, T_c, p
3. Compute B_a, B_b, B_c
4. If $R_{a,max} < p * T_a$

    $\delta T_a \longleftarrow \max(T_a * f,\ \Delta)$ (where e.g., $f = 0.0001$ Erlang, $\Delta = 0.0005$ Erlang)

    $\delta R_a \longleftarrow \delta T_a * (1 - B_b)$

    If $(T_a - \delta R_a) \geq (T_b + \frac{\delta R_a}{6})$

        $T_a \longleftarrow T_a - \delta R_a$

        $T_b \longleftarrow T_b + \frac{\delta R_a}{6}$

        $R_{a,max} \longleftarrow R_{a,max} + \delta T_a$

5. If $R_{b,max} < \frac{p}{2} * T_b$

    $\delta T_b \longleftarrow \max(T_b * f,\ \Delta)$ (where e.g., $f = 0.0001$ Erlang, $\Delta = 0.0005$ Erlang)

    $\delta R_b \longleftarrow \delta T_b * (1 - B_c)$

[2]This situation may arise when the initial traffic loads in tier B, and tier C cells are almost at the same level, and very low, whereas the traffic load in cell A is at a much higher level, and the fractional ARS coverage, p, has an intermediate value (For example, in Fig. 5, this occurs approximately when $0.16 < p < 0.38$).

If $(T_b - \delta R_b) \geq (T_c + \frac{\delta R_b}{2})$

$T_b \longleftarrow T_b - \delta R_b$

$T_c \longleftarrow T_c + \frac{\delta R_b}{2}$

$R_{b,max} \longleftarrow R_{b,max} + \delta T_b$

6. Recompute B_a, B_b, B_c
7. If $B_a = B_b = B_c$, then END
8. If $R_{a,max} \geq p * T_a$ and $R_{b,max} \geq \frac{p}{2} * T_b$, then END
9. If $B_a = B_b$ and $R_{b,max} \geq \frac{p}{2} * T_b$, then END
10. If $R_{a,max} \geq p * T_a$ and $B_b = B_c$, then END
11. Repeat Steps (4) to (10)
12. END

For numerical results, the parameter values considered are as follows. Number of wireless channels per cell (m) is 50. Originally, tier B cells are assumed to be operating at 2% call blocking probability, and tier C cells are operating with 1% call blocking probability (refer to Fig. 3). Corresponding traffic intensities in tier B and tier C cells are $T_b = 40.25$ *Erlang* and $T_c = 37.90$ *Erlang*, respectively. Traffic intensity in cell A is assumed to be increasing from $T_a = 40.25$ *Erlang* (corresponding to 2% blocking probability) onwards.

In the load balancing performance evaluation conducted in this paper, physical layer issues are not taken into account. In other words, our results are based on perfect physical channel condition. Also, issues like hand-off priority are not taken into consideration. Fraction of calls in a cell, covered by ARSs is p. With a single ARS it is assumed that $p = 0.04$ [3].

Fig. 4 shows the impact of ARS coverage on call blocking probability in cell A with relaying. We note that only partial ARS coverage is necessary to achieve proper load balance with relaying. However, the ARS coverage requirement for perfect load balancing depends on the "heat" level of the cell (cell A, in Fig. 3) with respect to the surrounding cells. This is because the "hotter" (i.e., more congested) the cell is, the more traffic needs to be relayed which, in turn, is dependent on the ARS coverage.

In Fig. 5 the impact of ARS coverage on all three tier cells (cell A, tier B and tier C cells) are plotted. As expected, with more ARS coverage, since the excess "heat" is distributed among the surrounding cells, call blocking probability increases in those cells (tier B, and tier C cells) while serving the extra traffic from the hot cell (cell A). Also, observe that, since tier B and tier C cells have very little difference in initial traffic, load balancing among them is achieved at much lower value of p.

[3]Typical values are : radius of cellular coverage is 1 *km*, and radius of a single ARS's coverage is 200 *m*. Hence, $p \approx \left(\frac{r}{R}\right)^2 = 0.04$.

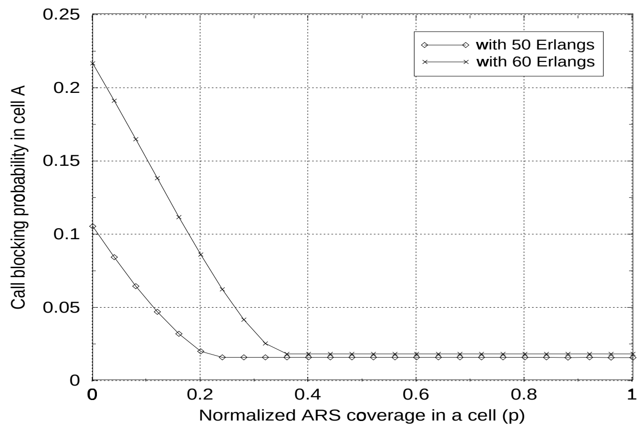

Figure 4: Call blocking probability in cell A versus normalized ARS coverage, with two different values of T_a. $T_b = 40.25$ *Erlang*, $T_c = 37.90$ *Erlang*, and $m = 50$.

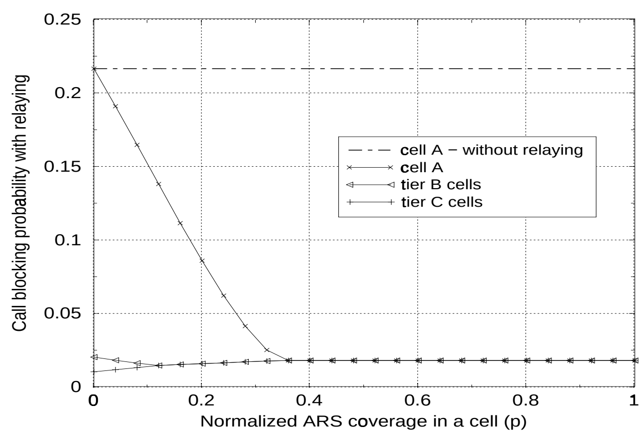

Figure 5: Call blocking probability versus ARS coverage in a cell. $T_a = 60$ *Erlang*, $T_b = 40.25$ *Erlang*, $T_c = 37.90$ *Erlang*, and $m = 50$.

In Figures 6 we show the effect of the two relaying schemes on call blocking probability with different traffic intensities in cell A. Observe that ARS coverage has a profound impact (up to 70%, in our example) on the relaying performance.

Traffic capacity enhancement in our 3-tier iCAR system model is shown in Table 1 and Fig. 7. Note that the capacity of iCAR system increases almost linearly, as the ARS coverage area in a cell increases. Observe from Table 1 that in our 3-tier cell model, the proposed iCAR system supports approximately 28 *Erlang* more traffic in a hot cell (cell A) (with approximate normalized ARS coverage of $p = 0.42$), for a given (2%) system-wide call blocking probability. In Fig. 7, the truncation of the curve beyond a point (approximately 28 *Erlang*) occurs because we restrict the acceptable

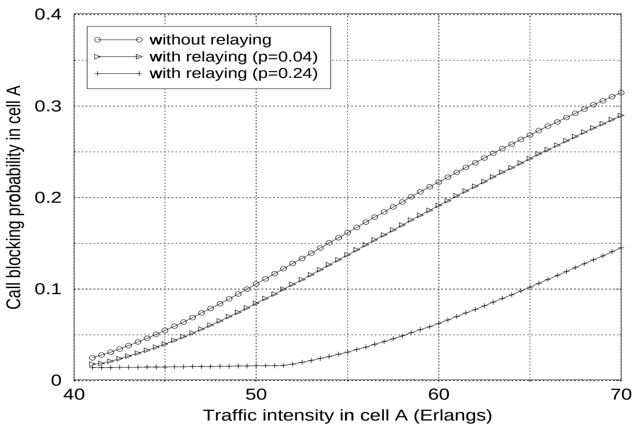

Figure 6: Call blocking probability versus traffic intensity in cell A with two different fractional ARS coverage. $T_b = 40.25$ *Erlang*, $T_c = 37.90$ *Erlang*, and $m = 50$.

Table 1: Performance of iCAR system for different normalized ARS coverage

Normalized ARS coverage (p)	Capacity increase (*Erlang*)
0.04	1.64
0.08	3.44
0.12	5.40
0.14	6.44
0.16	7.52
⋮	⋮
0.42	27.94

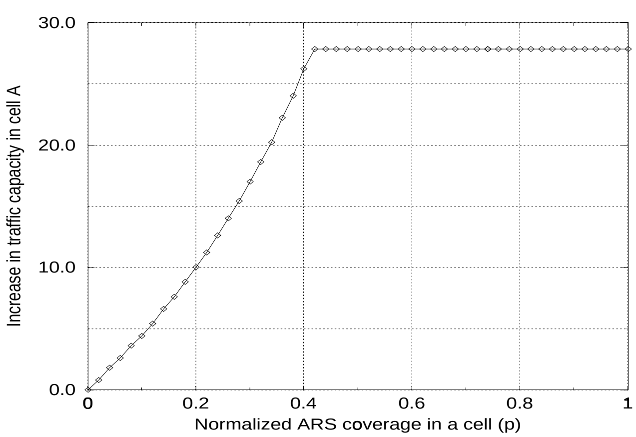

Figure 7: Capacity enhancement in iCAR system. $T_b = 40.25$ *Erlang*, $T_c = 37.90$ *Erlang*, and $m = 50$. T_a is increased from 40.25 *Erlang* onwards.

call blocking probability of the system to 2%.

5 Conclusions

In this paper, a detailed performance analysis of iCAR system - a novel next generation wireless network architecture - is presented. It is shown that traffic capacity in a cell increases almost linearly with ARS coverage. For a certain GoS, with sufficient ARS coverage traffic capacity enhancement is quite substantial. Depending on the possible amount of excess traffic (i.e., traffic imbalance) in a cell, the network operator can decide on deploying additional relay stations, thus keeping a balance between added system cost and necessary capacity enhancement.

It may be noted that the limitations (e.g., channel locking, limited area of coverage overlap) in conventional load balancing schemes without relaying technique do not exist in the iCAR system. However, because of introduction of ISM band channels as facilitators in channel borrowing, in addition to the cost factor, the problems associated with the ISM band communications (e.g., interference) are introduced. Also, ARS management and dynamic routing requirements associated with this scheme add to the performance overhead.

References

[1] http://www.ncs.gov/N2/Wps/cps1.html/.

[2] Y.D.Lin, and Y.C.Hsu, "Multihop cellular: A new architecture for wireless communication," in *Proc. IEEE INFOCOM'2000*, pp. 1273–1282, 2000.

[3] C. Qiao, H. Wu, and O. Tonguz, "Load balancing via relay in next generation wireless systems," in *Proc. of IEEE Mobile Ad Hoc Networking & Computing*, pp. 149–150, 2000.

[4] H. Wu, C. Qiao, and O. K. Tonguz, "A new generation wireless system with integrated cellular and mobile relaying technologies," in *Proc. of Wireless Access Systems Conference (WAS'2000)*, pp. 55–62, San Francisco, Dec. 2000.

[5] http://www.nwr.nokia.com/.

[6] T. Rappaport, *Wireless Communications: Principles and Practice.* Prentice Hall, 1996.

[7] V. Garg and J. Wilkes, *Wireless and Personal Communications Systems.* Prentice Hall, 1996.

[8] S. Maloo and C. Qiao, "Efficient routing and fast handoff in a mobile ATM network with a hybrid topology," in *5th Intl. Conf. on Info Systems Analysis and Synthesis (ISAS)*, vol. 4, pp. 614–621, 1999.

[9] S. Das, R. Castaneda, J. Yan, and R. Sengupta, "Comparative performance evaluation of routing protocols for mobile, ad hoc networks," in *7th Intl. Conf. on Computer Communications and Networks (IC3N)*, pp. 153–161, 1998.

[10] Y.-B. Ko and N.H.Vaidya, "Location-aided routing (LAR) in mobile ad hoc networks," in *ACM/IEEE the 4th Annual Intl. Conference on Mobile Computing and Networking (MobiCom'98)*, pp. 66–75, 1998.

Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS-35'02)
0-7695-1435-9/02 $17.00 © 2002 IEEE

[11] C. Toh, *Wireless ATM and Ad Hoc Networks: Protocols and Architectures.* Kluwer Academic Publishers, 1996.

[12] C. Perkins and P. Bhagwat, "Highly dynamic destination sequenced distance vector routing(dsdv) for mobile computers," in *Proc. of ACM SIGCOMM'94*, pp. 234–244, 1994.

[13] C. Perkins and E. Royer, "Ad hoc on demand distance vector routing," in *Proc. of IEEE WMCSA'99*, pp. 90–100, 1999.

[14] V. Park and M. Corson, "A highly adaptive distributed routing algorithm for mobile wireless networks," in *Proc. of IEEE INFOCOM'97*, pp. 1405–1413, 1997.

[15] D. Johnson and D. Maltz, "Dynamic source routing in ad hoc wireless networks," *Mobile Computing*, vol. 5, pp. 153–181, 1996.

[16] S. Murthy and J. Garcia-Luna-Aceves, "An efficient routing protocol for wireless networks," *ACM/Baltzer Mobile Networks and Applications*, vol. 1, no. 2, pp. 183–197, 1996.

[17] A. Iwata, C.-C. Chiang, G. Pei, M. Gerla, and T.-W. Chen, "Scalable routing strategies for ad hoc wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 8, pp. 1369–1379, 1999.

[18] S. Basagni, I. Chlamtac, V. Syrotiuk, and B. Woodward, "A distance routing effect algorithm for mobility (dream)," in *Proc. of the ACM/IEEE MobiCom '98 Conference*, pp. 76–84, 1998.

[19] G. Stuber, *Principles of Mobile Communication.* Kluwer Academic Publishers, 1996.

[20] R. Kohno, R. Meidan, and L. Milstein, "Spread spectrum access methods for wireless communications," *IEEE Communications Magazine*, vol. 33, no. 1, pp. 58–67, 1995.

0-7695-1435-9/02 $17.00 (c) 2002 IEEE

Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS-35'02)
0-7695-1435-9/02 $17.00 © 2002 IEEE