Inter-Operator Spectrum Sharing in a Broadband Cellular Network

Gareth Middleton*, Kari Hooli†, Antti Tölli†, Jorma Lilleberg*
*Nokia Technology Platforms
Oulu, Finland
Email: gbmidd@ee.oulu.fi
†Centre for Wireless Communications
University of Oulu
Oulu, Finland
Email: kari.hooli@ee.oulu.fi

***Abstract*— Inter-operator spectrum sharing in a wide-area broadband network is considered in this paper. A packet-based cellular model is developed, emphasizing the shift in the telecommunications industry towards IP-based services. In the framework of this architecture, we show the ideal maximum load that the network can handle without excessive delays, and present our main result: that even with sharing, this operating point is rarely achieved. This is the consequence of geographical and physical layer constraints, which limit the gains achievable by spectrum sharing. Using a physical layer cellular model with idealistic resource management, we quantify the achievable sharing gains. We further analyze these gains in the context of variable data rates, quality of service guarantees, and number of operators.**

I. Introduction

As mobile communication terminals continue to become more widespread and the data rates supported by their networks increase, there will exist a need to understand how to extract more capacity from one of the most limited resources: the spectrum. The most obvious solution involves the licensure of more bandwidth, but this is not necessarily the most efficient use of the resource. Rather, operators must turn to each other to utilize the spectrum available to them with higher efficiency [1]. This issue has already been examined from a policy-standpoint [2], and also from a small-scale physical perspective [3]. Some work exists on the large scale gains [4], but there are few results on the impact of sharing on a wide area packet network with attention paid to the physical layer [5].

In this paper, we will construct a framework within which to analyze sharing gains in a packet-based cellular network. We choose a packet network to reflect the increasing demand for Voice-over-IP (VoIP) and broadband data services [6]. We will investigate how much spectrum we can save by using sharing techniques, while still preserving the quality of service (QoS) for all users. In this paper, we will focus on how to preserve QoS for these applications while intensifying spectrum use and, therefore, providing service to more subscribers.

Several issues must be considered in such an analysis. Most importantly, the operating point of the simulated system must be determined–at what level of load does the system require sharing algorithms to preserve QoS? We will attempt to exploit the differences in traffic profiles, releasing bandwidth from an off-peak operator to be used by an overloaded operator. We will show that the maximum level of load is upper bounded by the ideal share case, and then quantify the difference between the gains under ideal sharing conditions and the gains achievable in a more realistic physical environment. Further, we will also quantify the gains from any sharing over the case in which no sharing occurs. Our main result concerns the differences in sharing gains as the physical layer and geographical constraints are taken into account.

II. Proposed Network Architecture

To make sharing feasible, we must assume that all operators use the same radio access technologies. In this case, the radio access model reduces from a more complicated frequency division-based sharing among operators to a time-division duplex (TDD) system, as is being discussed for 4G implementations [7]. Each operator is allocated a certain number of slots in a super frame for both uplink and downlink. Sharing spectrum therefore amounts to operators trading timeslots, and empty timeslots within each operator's allocation repre-

TABLE I

TYPES OF SERVICE, DATA RATE DEMANDS, AND DURATIONS

Service	Data Rate	Mean Duration
Voice	19.2 Kbps	180 sec
Web	400 Kbps	5 sec
Photo Message	1 Mbps	15 sec
Interactive Video	1 Mbps	$>$ 600 sec
Streaming Video	2 Mbps	30 sec

TABLE II

RELEVANT PHYSICAL LAYER PARAMETERS

Parameter	Value
Modulation	16-QAM
Bandwidth	20 MHz
FFT Size	256 + 64 cyc. pref.
No. Data Carriers	192
Noise Figure	9 dB
Super Frame Duration	10 ms
Minimum Allocation entity	OFDM symbol

sent bandwidth available to other operators [8]. Our TDD system will use orthogonal frequency division modulation (OFDM), with a 16-QAM constellation providing maximum downlink data rates of 46 Mbps to a single user in ideal conditions (see II for physical parameters).

We will consider a wide area network, spread over approximately 1.5 km^2. Operators will not use collocated base stations, though each will ensure that it has coverage over the entire region. Subscriber units will be randomly placed within this area, and will be randomly assigned to a home operator from whom they regularly buy service. Each subscriber will also be assigned a service class, ranging from voice to internet-video (Table I), to give a variety of data rate demands. We will consider only downlink sharing, as this is where peak traffic is most likely to occur.

We will use a standard Poisson traffic model for each class of service, generating downlink data requests at specified rates [9]. The length of calls will be drawn from an exponential distribution, with means determined according to the type of service [10].

The physical layer parameters of our model are summarized in table II.

Operators will use their slot allocations to schedule data transfers to their mobiles, and the sharing algorithm will attempt to schedule any packets that cannot be accommodated by their home operators. The algorithm must account for geographical location of time slot allocation, as slots in use in certain sectors of the region may be free in other sectors. In currently deployed systems, this is the issue of flexible carrier allocation schemes.

It should be noted that in this network simulation, we have assumed that interference between mobiles does not occur. We idealize the time synchronization among operators so that the slots never overlap, and time slot reuse is perfect over the sectors.

A. *Ideal Results*

If we assume ideal conditions, we can state the maximum load under which the system can operate. Clearly, this will occur when all slots are full. If we assume s slots and a capacity of b bits per slot, then the system can send sb bits per downlink super frame per sector. It has been shown in [4] that ideal capacity increases asymptotically as the number of operators increases, to be limited finally by the capacity of the base stations. Here, this corresponds to using the full 46 Mbps transfer capacity in each cell.

B. *Scheduling-Sharing Algorithm Outline*

To approach this ideal capacity, we define a scheduling-sharing algorithm to fill the super frame with packet data. We assume that operators share resources only as a “last resort.” To incorporate this constraint into the algorithm, we allow operators to have a dedicated set of slots in the super frame, which are allocated only to subscribers of that operator.

Because we are focusing on physical layer effects, we assume an idealistic scheduler which is quite centralized and requires a high degree of signalling between operators. This scheduler iterates through the base stations in a geographic order and assigns timeslots to the mobiles, maximizing slot reuse, and finally checking for overloaded operators. If a particular mobile lacks service and an empty slot exists in another operator’s allocation within range of the mobile, the scheduler arranges a share and assigns the mobile to the other operator. In the case that more than one operator has excess capacity, the scheduler uses a round-robin scheme to assign the shares [4].

If no excess capacity exists, the unscheduled packets experience a delay in service and take priority in the next super frame. In an entirely overloaded system, delays are on the order of several super frames and QoS is degraded.

III. Performance Metrics

Up to now, we have been discussing the importance of Quality of Service as it relates to packet-switched networks, and we have been emphasizing the need to preserve it in overloaded systems. In order to measure QoS, we will consider downlink delay time. For a streaming application, data arrive from the IP network at the base-station buffer at some nominal rate, and wait for transmission to the mobile over the air interface, ideally at the same rate.

Buffering incoming data at the mobile can smooth downlink delays, but for practical reasons, this buffer cannot be too large. As such, one of our Quality of Service measures becomes the amount of data remaining in the downlink buffer at the base station. We will measure this data in frames, each of which is one unit of coded data ready for transmission. In the analyses to follow, the frame is the fundamental unit of data.

We will also consider *service interruptions*, which occur when the buffer at the mobile empties. For this to occur, the buffer at the base station must contain at least as much data as the nominal capacity of the mobile buffer. If the mobile buffer empties, the user experiences a pause in the data stream and QoS is degraded. We model the system as dumping the data at the base station in this event, such that the untransmitted information is discarded and never reaches the mobile. The stream resumes after the buffer has been cleared, with the possibility that network conditions have improved. In this way, large backlogs of data do not accumulate in the downlink buffer at congested base stations. The size of the mobile buffer determines how much data may become backlogged due to congestion, and adjusting this parameter changes overall network delay while also affecting the number of service interruptions. Initally, we assume a mobile call buffer capable of buffering up to 1.5 seconds of data for each user, regardless of service class.

IV. Simulation Results

Fig. 1 is a map of our tower placement in a dual operator scenario, which was used for the majority of simulation. Notice that only the center tower coincides with that of the other operator; otherwise there is no co-location and the sectoring angles are offset.

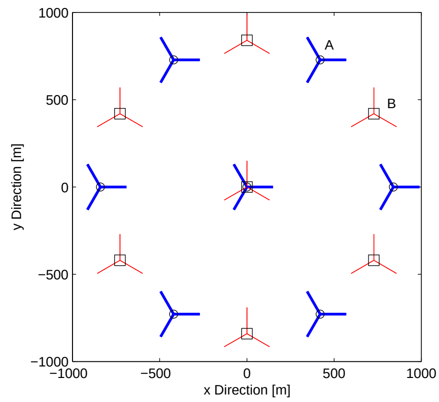

Fig. 1. The location of base stations is shown here, with sectoring divisions, for the dual operator case. Towers belonging to operator A are shown in blue, those belonging to operator B are in red.

TABLE III
System-wide Operating Point

Service Class	Average No. of Users	% Tot. Data
Voice	1600	2.5
Web	300	9.6
Photo Msg	200	16.0
I-Video	100	8.0
Streaming Video	400	63.9

A. Sharing Gains and Ideal Throughput Capacity

1) Ideal Throughput: The parameters given in table II lead to a superframe lasting for 10 ms and consisting of 100 data frames, each of which can carry 1536 bits of uncoded information. Each of the three sectors on a base station has a full superframe available, and we have modeled 14 base stations in the dual operator scenario. Maximum network throughput is calculated as the product of all of these values, which comes to about 644 Mbps over the region, both operators combined.

The system can be made to operate at maximum throughput capacity if enough load is applied, but in these cases, we also observe high levels of service denials.

2) Operating Point: The *operating point* of the system is the average level of downlink load faced by the network relative to its capacity. We would like to find an operating point at which the system experiences delays without sharing, but when sharing becomes available, the system is much more able to handle the same amount of load. It as at such an operating point that we show our main result: that although there is capacity for all calls

TABLE IV
AVERAGE FRAME DELAYS FOR EACH CLASS OF SERVICE

Service Class	**Non-Shared**	**Shared**	**% Decrease**
Voice	0	0	0
Web	0	0	0
Photo Msg	102	14	88
I-Video	486	62	91.5
Streaming Video	5711	1016	86.2
Avg Excess Frames	1850	1786	3.4

to be handled on the shared spectrum, delays still exist.

Since load is affected by data rate demands, call duration, and number of mobiles, the set of operating points satisfying this requirement is very large. Empirically, we found an interesting point to be defined as in III while using the rates given in I. In the dual operator scenario discussed below, this is the level of load considered.

3) Maximum Load Capacity Results: In our simulation, we monitored data frames which became backed up in the downlink buffer, in both the shared and unshared cases. Taking an average over 50 unique mobile distributions and 2-minute call-demand profiles, the differences between the shared and unshared case at our chosen operating point are given in IV. These averages represent the mean number of delayed frames in the system, according to service class.

It can be seen in IV that low data rate and low duration services (Voice and Web) do not experience any delays in either case. This is a consequence of the round-robin packet scheduler, which provides service to all classes equally until the frame is full, meaning that the low requirements of Voice and Web services are always met. There is a dramatic decrease in delayed frames for higher demand services when sharing is allowed, showing that sharing provides substantial gains at this operating point. This happens because neighboring sectors provided by the other operator often have enough space to accommodate excess traffic. Even though both operators are experiencing high load, allowing the system to exploit temporal shares provides significant gains.

4) Main Result–Maximum Capacity is Not Achieved: Notice in IV that the number of average excess frames does not decrease significantly as sharing is introduced. The average excess frames is a measure of how many frames are between the network capacity and its utilization, averaged over all super frames. The result that it changes little between the shared and unshared case suggests that the gains from sharing are temporal in nature and are not caught by the time average.

Our key result: that even with sharing available, the system still experiences delays and further, *there are more excess frames than delayed frames.* If the system was ideal, there would be zero delays in this case, since enough excess capacity exists to handle the delayed data. The reason this doesn't happen relates to the spatial distribution of mobiles: cells with high excess capacity do not necessarily have excess demand nearby and vice versa. Further, channel conditions may allow a lower data rate on a shared channel than a home channel, such that sharing is not able to fully compensate for the delay.

B. Quality of Service

QoS is also well-measured in terms of service interruptions, which occur when the mobile buffer empties as described previously. A smaller mobile buffer should result in a greater number of service interruptions, since less congestion is tolerated, *ceteris paribus*. However, congestion in the downlink buffer caused by a larger mobile buffer can persist for a longer period of time, before being cleared either by transmission or data dumping. Thus an interesting tradeoff exists between mobile buffer size and overall network congestion, as shown in Fig. 2: smaller mobile buffers experience more interruptions, but larger mobile buffers cause more network congestion in our model.

We would like to evaluate the QoS issue in both the shared and non-shared cases. From the plot, it can be seen that gains in terms of interruptions are largest at lower levels of QoS, but that network congestion is more dramatically affected by sharing at higher QoS levels.

This expands our result from the previous section: that sharing affects different types of service in different ways. In the service interruption metric, sharing helps least for long duration QoS delay because the only services vulnerable to that kind of interruption are high rate, long duration services such as streaming video. Shorter, lighter loads such as voice are better able to use buffering and so do not suffer interruption with larger buffer–hence the smaller gain from sharing.

Similarly, the gain from sharing in terms of Frame Delay is most dramatic for larger buffer size (higher QoS delay time) because the network can flexibly share spectrum to prevent packets from accumulating. At smaller buffer sizes, even a small delay amounts to an inter-

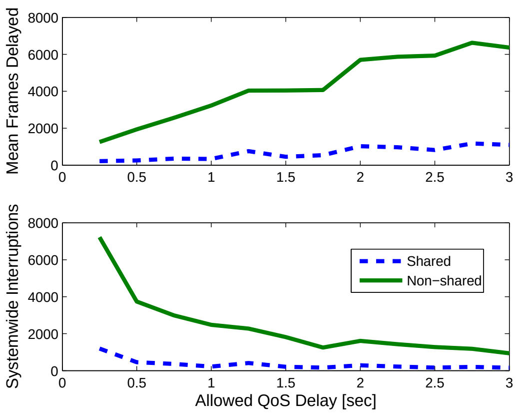

Fig. 2. Simulation Results: As QoS delay time increases, gains from sharing decrease, as measured by total number of service interruptions yet increase in terms of system-wide average frame delay.

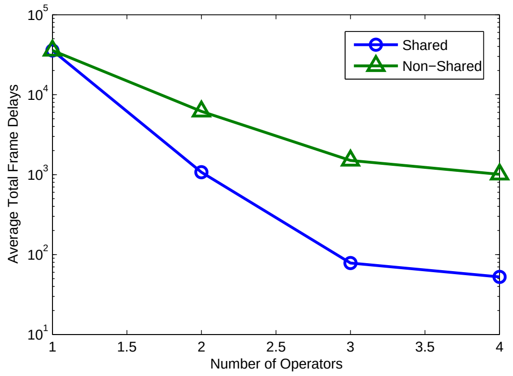

Fig. 3. Simulation Results: As Number of Operators increases, the gains from sharing increasing also. Demand scales with operators, but unshared delays decrease because with more mobiles, the random distribution of loads allows for more empty slots.

ruption, which triggers a data dump–this keeps overall delay low, but interruptions high. Sharing allows the delayed packets to be transmitted, dramatically lowering the interruption count but doing little for the already low delay average.

C. Number of Operators

A more realistic simulation would increase the number of operators. To introduce this degree of freedom to our simulations, we modified the map given previously (Fig. 1) by adding more hexagonally arranged towers belonging to different operators, offset geographically from those already present such that towers are not co-located. We assume the same operating point as before, accordingly scaling the call load with the additional operators. According to [4], the average network throughput should increase with the number of operators if sharing is allowed. This corresponds to a decrease in the number of delayed frames for all classes of service.

Fig. 3 shows the average frame delay for the cases of one to four operators. The gains from sharing become larger as the number of operators grows, because there are more opportunities for a share to be arranged. This is seen in Fig. 3. If we increase the number of operators further, average total frame delay becomes asymptotically bounded by the operating point of the network.

V. Conclusions

These results have shown that sharing the spectrum provides an increase in overall spectral efficiency, as viewed from a variety of standpoints.

Our key result is that sharing provides a roughly 85% performance improvement over conventional fixed spectrum allocations, although ideal sharing gains cannot be achieved among systems with non-co-located base stations. Hence, we must look to the development of sharing algorithms to maximize realizable gains. We must find methods exploit gaps in service demand, while using as little of the backbone network as possible.

References

[1] T. A. Weiss and F. K. Jondral, "Spectrum pooling: An innovative strategy for the enhancement of spectrum efficiency," *IEEE Commun. Mag.*, vol. 42, pp. s8–s14, Mar. 2004.

[2] J. M. Peha, "Spectrum management policy options," *IEEE Commun. Surveys*, Dec. 1998.

[3] A. Pezeshk and S. A. Zekavat, "Ds-cdma vs. ms-cdma, a performance survey inter-vendor spectrum sharing environment," in *Conf. Rec. of 37th Asilomar Conf. on Signals, Systems and Computers*, Asilomar, California, 2003, pp. 459–464.

[4] B. Aazhang, J. Lilleberg, and G. B. Middleton, "Spectrum sharing in a cellular network," in *8th Annual IEEE Symposium on Spread Spectrum Applications and Technologies*, 2004, pp. 355–359.

[5] T. A. Weiss, M. Spiering, and F. K. Jondral, "Quality of service in spectrum pooling systems," in *IEEE International Symposium on Personal, Indoors, and Mobile Radio Communications*, vol. 1, Sept. 2004, pp. 345–349.

[6] V. G. Cerf, "On the evolution of internet technologies," *Proc. of IEEE*, vol. 92, pp. 1360–1370, Sept. 2004.

[7] R. Esmailzadeh, M. Nakagawa, and A. Jones, "Tdd-cdma for the 4th generation of wireless networks," *IEEE Wireless Commun. Mag.*, vol. 10, pp. 8–15, Aug. 2003.

[8] P. Leaves, M. Breveglieri, C. Hamacher, D. Grandblaise, F. Migneret, K. Moessner, and D. Bourse, "Dynamic spectrum allocation and system coexistence in reconfigurable multi-radio networks," in *IST Mobile Summit*, Aveino, Portugal, 2003.

[9] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems*, 2nd ed. New York: Kluwer Academic/Plenum Publishers, 2000, p. 346.

[10] H. Nyberg, C. Johansson, and B. Olin, "A streaming video traffic model for the mobile access network," in *VTC 2001 Fall*, 2001, pp. 423–427.