CAPACITY GAINS THROUGH INTER-OPERATOR RESOURCE SHARING IN A CELLULAR NETWORK

Ville Heinonen, Pekka Pirinen, and Jari Iinatti
Centre for Wireless Communications, University of Oulu
Oulu, Finland

ABSTRACT
Due to a growth of the amount of mobile phones and, at the same time, increased data rates, the need for radio resources is proliferated. Since the spectrum is a finite resource, allocated channels must be used as efficiently as possible. Nevertheless, efficient resource utilization in cellular networks is complicated because loading of networks varies over time and place. The aim of this paper is to exploit differences in the loadings via resource sharing between operators. Since the operators will cooperate, it is possible to use resources more efficiently, and on the other hand, a temporary lack of resources may be avoided. We evaluate three different sharing cases, and moreover, compare these with the conventional non-sharing case in order to find out achievable sharing gains. The studied sharing cases are "sharing as a last resort", "always connected to the least loaded" and "sharing as a secondary user". Finally, we analyze the achievable sharing gains in terms of service probability and compare different implementation alternatives. According to the results, notable sharing gains can be attained with each sharing algorithm and hence, spectrum efficiency is increased. The best performance is achieved with the always connected to the least loaded case. Furthermore, services with moderate data rates are the most suitable for resource sharing.

I. INTRODUCTION

The use of mobile communication terminals has grown at a tremendous rate during the past two decades. At the same time, bit rates have increased due to the use of multimedia applications, which inevitably translates into a large spectrum demand. The above, coupled with maintaining a sufficient coverage area, is leading towards a scarcity of spectrum. These issues have increased the demands for operators to provide an ubiquitous network access to all clients. Today's wireless networks are regulated by a fixed spectrum assignment policy. Currently, the spectrum is divided into a set of disjoint blocks, which are assigned to different individual operators on a long-term basis. Thus, every operator is assigned a proprietary fixed band for exclusive use, and interference between operators is minimized. Consequently, the overall spectrum utilization is poor. Traffic in cellular systems is heterogeneous with spatial and diurnal fluctuations, which causes problems in efficient resource allocation. According to the Federal Communications Commission (FCC), a large portion of the assigned spectrum is used occasionally, with utilization varying in the range of 15% to 85% in the bands below 3 GHz [1]. Under these circumstances, a particular operator may face a situation with a lack of resources, while there may be a huge amount of unused resources in another operator's network. This problem can be alleviated by resource sharing among competing operators. In addition, there are also economical aspects involved. It is important to utilize the spectrum efficiently in order to reduce the number of base stations needed and hence costs. All this leads to migration from fixed allocation to more flexible resource management.

Flexible spectrum management has been widely considered a "hot topic" and a number of different working groups, e.g., WINNER [2], are involved in efforts to tackle this issue. However, resource sharing among operators is a fairly new issue, only a few related papers exist. For instance, papers [3], [4] and [5] study resource sharing among operators and evaluate the results in terms of average frame delays and average number of drops. Moreover, the case where operators share their resources only as a "last resort" is studied in these papers. Furthermore, papers [4] and [3] take into account the case where mobile stations (MS) are always connected to the closest base station (BS), whether it belongs to its home operator or not.

Sharing in this context is defined as co-operation of two or more operators in the process of building, maintaining and upgrading a mobile network. Each of the operators has its own customer base, but in some cases, such as being out of resources, one operator may need to turn to an other operator and use its nodes to gain access to the mobile network [6].

In this paper, different sharing algorithms are studied and the achievable sharing gains are quantified. Resource sharing between operators is based on the exploitation of differences between the loadings of networks. Three different sharing cases are evaluated. First one is "sharing as a last resort" case, which is very straightforward method. Mobile stations enquire access to other operators' networks only if the home operator is out of resources. Hence, resources from the home operator are always used first, and sharing occurs only as a last resort. Second sharing case is "always connected to the least loaded" that is based on full co-operation among operators and is in a way opportunistic. Each mobile station is always connected to the base station that has the greatest amount of unused resources, whether it is its home operator or not. Lastly, "sharing as a secondary user" is evaluated, which differs from the other cases because it does not have quality of service (QoS) guarantees for all services due to different prioritization among users of home and foreign operators. Users of the home base station are referred to as primary users, whereas the users of foreign base stations are considered as secondary users. Hence, primary operator has always priority to recall the shared resources to its own use and hence does not sacrifice its own performance. By this way secondary system may use unused resources as long as they are available. Finally, sharing cases are compared with
case where mobile stations are only assigned to their home operator. This case is referred to as non-sharing case. Sufficient amount of simulations have to be run in order to achieve reliable and comparable results. After completion of the simulation runs the results are averaged and compared in terms of service probability. Service probability is defined as a probability that the user is served by the system in any location within the area. Thus, greater service probability means better system performance. We have defined the service probability as a ratio of successfully served services and the total amount of terminated services. When the service is terminated, it is either successfully served or interrupted.

This paper is outlined as follows. Section II depicts the system model used in the simulations. Section III focuses on the simulations and the main assumptions. Key parameters are also explained. The simulation results are addressed in Section IV. In addition, different results are evaluated, and sharing gains are quantified. Finally, conclusions are presented in Section V.

II. SYSTEM MODEL

The system in question is based on a packet-switched cellular network that operates under a time-division scheduled medium access scheme. Hence, the channel is divided into a sequence of frames that are further divided into fixed-size portions, referred to as time slots. The operators share their slots among the users for data transmission. Thus, the resource sharing is basically done by sharing the time slots. A simple fixed channel assignment scheme is used, and so every BS has an equal amount of slots. Figure 1 illustrates the simulation area that is approximately 4.0 km$^2$ and the entire simulation region can be served by all competing operators.

As can be seen, four competing operators are located in the same geographical area and each operator has seven base stations within the simulation grid. Every base station covers three sectors, and there is a full superframe available in each sector.

![Figure 1: Simulation scenario.](image)

Table 1: Traffic parameters for different service classes

<table>
<thead>
<tr>
<th>Service</th>
<th>Mean duration</th>
<th>Data rate</th>
<th>Amount of users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice</td>
<td>120 s</td>
<td>19.2 kbps</td>
<td>4200</td>
</tr>
<tr>
<td>Web</td>
<td>5 s</td>
<td>400 kbps</td>
<td>1050</td>
</tr>
<tr>
<td>File transfer</td>
<td>13 s</td>
<td>1 Mbps</td>
<td>1050</td>
</tr>
<tr>
<td>Video Stream</td>
<td>45 s</td>
<td>2 Mbps</td>
<td>700</td>
</tr>
</tbody>
</table>

Although the placement of BSs is highly idealistic, the variability of the data link statistic is ensured by BS antenna offsets and spatially uniformly distributed mobile station locations.

All operators have equal averaged traffic demand and thus the worst case scenario is considered. Hence, it is obvious that if a certain operator is highly loaded and out of resources, it is likely that the load of other operators’ BSs is also high. However, traffic in the base stations varies among base stations due to the stochastic nature of incoming traffic. Traffic in the network under consideration is heterogeneous, so different service classes are taken into account. Moreover, different services have class-specific mean durations, data rates and number of users, and thus different requirements from the network. These parameters are tabulated in Table 1. A service request arrival process is modeled by using a Poisson random generator, whereas the duration of each service class follows the exponential distribution [4].

According to delay requirements, service classes can be further categorized into two types of services: real-time (RT) and non-real-time (NRT). Voice and web classes are considered as real-time services. By the contrast, file transfer and video services are non-real-time services.

Each MS can carry only one connection at a time. The number of users of different service classes is predefined, but several arrival rates are simulated in order to study the effect of loading. Then, the ratio between arrival rates of different classes is kept fixed.

III. SIMULATIONS

The simulation consists of 50 runs, where the users are randomly uniformly distributed within a geographical area and also randomly assigned to different operators in the beginning of each run. Thus, our results correspond to the cellular network, and simulation complexity is kept sensible. The simulation time is set to 600 seconds in order to ensure a sufficient amount of served services and other results. Slot allocations are made in every superframe. Several call patterns are simulated with different level of loading in order to generate loading curves. To ensure comparable results, same call patterns are
used in simulations of each sharing case. Simulation parameters are gathered into Table 2.

When the mobile station needs to be served, it sends a request to the home base station. The home BS makes a decision which base station will serve. This depends on the used sharing algorithm and loading conditions. If the target base station has not sufficient amount of resources, the service is blocked by the admission control. In that case, the user does not try to reconnect within simulation time. Thereby, call patterns remain similar in all cases. Thus, overloaded network can be avoided. In secondary user case, services may also be dropped due to exceeded delay constraints. For real-time services the delay constraint is set to 100 ms, whereas for the non-real-time services it is 2 s. The simulation metrics are set from the user point of view. The number of successfully served and interrupted services within the simulation time are calculated. The service probabilities are calculated after simulation of each drop. Finally, the results are averaged over different drops.

To have realistic simulation models, the physical layer and geographical constraints are taken into account. Simulation environment is urban, so the path losses are calculated using the Okumura-Hata model. Hence, the signal strength can be obtained because the transmission powers and antenna gains are known. Possible serving BSs are selected according to signal strength. Several assumptions have to be made to facilitate the complexity of the simulator, such as channel conditions and the locations of MSs are assumed to be static over one simulation run. Moreover, coverage areas of BSs do not overlap within one operator’s coverage range. Hence, BS with the greatest signal strength of each operator is considered as a possible BS to serve. In addition, services can be served by one BS at a time, and, serving BS will not change even if the loading conditions change dramatically.

IV. RESULTS

It is obvious that sharing gains depend on loading, and significant sharing gains can be attained only at high loads. Thereby, a high-loaded network is described, and several snapshots with different values of arrival rates are simulated in order to generate the loading curves. We are also interested in the performance of different QoS classes because they have different resource demands. Therefore, not only total gains are considered, but also class-specific sharing gains. Under these circumstances, it is possible to find out if different classes have unequal sharing gains, and moreover, if a certain class is more suitable for sharing than the other classes.

First, the different QoS classes are not distinguished, and thereby the total traffic with 7000 users is considered. Figure 2 illustrates the service probability versus the average number of service requests per second.

The different curves represent the different sharing algorithms. As can be seen, the always connected to the least loaded case outperforms the other cases. Therefore, by using that algorithm, the greatest number of services can be served without interruptions. However, in low and moderate values of arrival rates, differences between "always connected" and "last resort" cases are unsubstantial. The performance of the "secondary user" case is the worst in comparison with the other sharing cases. Despite of this, great sharing gains can be achieved, and resources are used more efficiently. The operators have the same amount of resources and equal traffic demands, thus it can be assumed that similar gains are obtained by each operator.

After the evaluation of total traffic, we turn our attention to services of different QoS classes. First, we take the voice service into examination. In Figure 3, the service probability versus the arrival rate is depicted. We can detect that sharing gains can be obtained. However, voice service has low resource requirements so that good level of service probability can be provided even without sharing. The service probability in non-sharing is over 99% even at high values of arrival rate. Thus, we can say that obtainable sharing gains for voice service are minor.

Service probability curves for web services are illustrated in Figure 4. Evidently, the resource sharing improves service probability, especially at high arrival rates. For instance, if we check the point where the arrival rate is 20 service requests per second, the service probability in all sharing cases is over 99%, while the corresponding value for the non-sharing case is about 12 percentage units worse. Significant sharing gains are attained even at low values of the arrival rate.

The services considered so far have had only small resource requirements, which are nearly always met when resources are shared between the operators. Hence, good level of service probability can be provided when sharing is allowed. Performance of the file transfer service is evaluated next. Service probabilities are depicted in Figure 5. We can see that a 100% service probability can be obtained at low values of the arrival rate regardless of the sharing algorithm, and in higher values of data rates, service probability decreases even with sharing. However, remarkable enhancement in terms of service probability can be attained via sharing. In addition, differences in performance of different sharing cases obviously exist. Again,
the best performance can be achieved with the always connected algorithm, and the difference between sharing as a last resort and secondary user is very small.

Lastly, video stream service is examined. It can be considered the most challenging service because it has a high data rate besides having a moderate mean duration. Service probabilities are addressed in Figure 6. As can be seen, the shapes of video stream curves differ from the former cases, and differences between cases shrink at high arrival rates. In addition, the service probability decreases quickly with the arrival rate. However, attaining notable sharing gains is possible at low and moderate values of arrival rates. Setting the sharing cases in an order of superiority is not unambiguous because it depends on the arrival rate. According to these results, the always connected case is the best option at low arrival rates, whereas its performance is the worst at high values of arrival rate. This is a consequence of a lack of resources when the resources have already been allocated to more active users of a different service class.

V. DISCUSSION

A lot of other open issues worth investigation have emerged during this research. Firstly, improving the present simulator and extending the performance metrics to include delay, throughput and spectrum efficiency are issues to be looked into. In addition, for instance, the signal strengths should be further considered, and the effect of adaptive modulation and coding (AMC) should be investigated. In the AMC scheme, the modulation method/order is changed according to the signal level. Hence, further optimization of the selection of the BS can be done according to the signal strengths. Therefore, resources are used even more efficiently, and performance of the system is further improved. So far we have evaluated the performance from the user point of view and different operators are not dis-
distinguished. The results would be studied also from the operator point of view and study the effect of the number of co-operative operators.

VI. CONCLUSIONS

The results of this research have proven that resources can be used much more efficiently via resource sharing among competing operators. Since operators share their resources, service probability can be increased significantly and a much greater number of users can be served without interruption. Notable sharing gains can be attained with each sharing algorithm, but always connected to the least loaded sharing case provides the best performance. It is also obvious that achievable sharing gains depend on the loading of network, but still remarkable sharing gains is attained even at low and moderate level of loading. In addition, certain QoS classes seem to be more suitable for sharing than others. It is obvious that if the service has low resources requirements, it can be served without sharing. On the other hand, when the service has high resource requirements, it may be blocked even with sharing due to an insufficient amount of resources. Thus, we can assume that a service with a moderate data rate is the most suitable for resource sharing. The results have demonstrated that the greatest sharing gains can be attained in web and file transfer services. Consequently, the services data rate has a significant influence on the attainable gain improvement. In summary, we can state that when resources are shared among operators, spectrum efficiency is improved, and hence, a better quality of service can be provided for customers. Therefore, both the customers and the service providers benefit from resource sharing.

ACKNOWLEDGMENT

The authors would like to thank the Finnish Funding Agency for Technology and Innovation (Tekes), Nokia, Nokia Siemens Networks, and Elektrobit for the project funding enabling this research.

REFERENCES