Inter-layer Resource Management for Hierarchical Cell Structures

Guevara Noubir

Data Communication Group, CSEM., CH-2007, Switzerland.
guevara.noubir@csemne.ch

ABSTRACT

Hierarchical Cell Structures (HCS) provide flexibility in handling non-homogenous traffic, but raises several problems to be efficiently deployed. In this paper, we show when and how resources can be shared between layers (e.g., TDMA frame partitioning between layers). We also provide several algorithms for inter-layer resource management (handover, admission control, congestion/load control).

I. INTRODUCTION

The aim of HCS is to handle non-homogeneous traffic in a flexible way (local hot spot traffic, service differentiation, increased traffic efficiency and VPN/CPN support). The cells in a given layer are more suitable to handle the traffic of some of the mobile stations (MS). By order of preference, a cell will service: 1) MSs that are not covered by any cell of lower layer. The primary objective of macro-cells is to serve a wide area while leaving to micro-cells the handling of local hot spot traffic. 2) Fast moving MSs that generates high number of Handovers (HO). This is because, HO induce a significant signaling traffic, which may degrade the network quality of service. 3) MSs covered by overloaded lower layer cells. This is to reduce the average blocking probability. MSs are ordered in this way because it is more “acceptable” for a high speed user to have its call blocked (or dropped) than it is acceptable for a mobile which is not moving.

On the other hand a MS can have a preference list of layers (or base stations (BS)) for various reasons (e.g., a MS will prefer to use a pico-layer cell when it is in its company’s private network (for obvious reasons of tariff, privacy,…), a fast MS would immediately prefer a cell layer that corresponds to it’s speed (layer-cell crossing average time)).
However, the introduction of HCS presents several problems that have to be solved to achieve an efficient implementation. These problems are related to the separation of layers in time or frequency and to the adaptation of resource management algorithms (handover, admission control, congestion/load control). Ignoring these problems may result in degrading the capacity of the overall system instead of improving it. To illustrate the different aspects of HCS implementation we will take examples from the FRAMES (Future Radio Wideband Multiple Access System) project partly funded by the European commission. The goal of this project was to define a proposal for UMTS (Universal Mobile Telecommunications System) radio access scheme. FRAMES proposed two radio schemes FMA1 (TDMA with and without spreading [1]) and FMA2 (wideband CDMA).

The paper is organised as follows: In section II, we will present the constraints on separating the layers and schemes for achieving an efficient resource sharing. In section III, we will discuss the implications of HCS on resource management algorithms and provide some algorithms to achieve efficient HCS implementations. Finally, in section IV, we will give some conclusion and directions for future research.

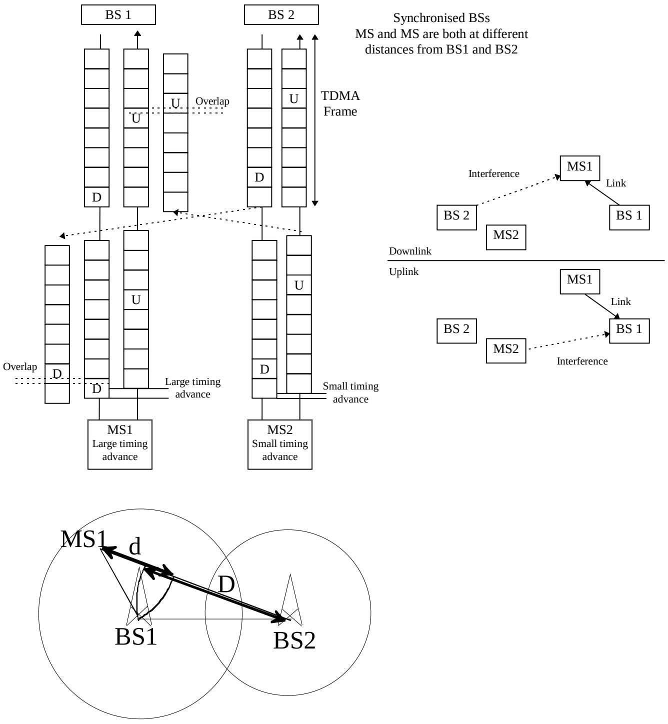
II. CONSTRAINTS ON LAYERS SEPARATION

There is potentially three ways of layer separation: code, time and frequency. It is obvious that it is not possible to separate two layers only by code while using the same frequency band. This is because MS connected to higher layer cells require much higher transmission power than lower cells. On the other hand separating the layers on a frequency basis by splitting the overall spectrum is possible [2]. However, a separation in frequency is more rigid than a separation in time. This rigidity results in less efficiency. In the following we investigate the possibility of two layers sharing a frequency by partitioning the set of available time slots. To achieve such goal it is necessary to synchronize the BSs. Synchronization is necessary but not sufficient. We assume to have a mixed environment of macro-cells and micro-cells (HCS of two-layers). Generalizing to more layers can be done by considering at each time a couple of adjacent layers (frequencies are shared only between adjacent layers).

A. INTERFERENCE DUE TO INDEPENDENT DELAY FROM MOBILES TO DISTANT BSS

Figure 1. Shows that even when the BSs are synchronized and two mobiles MS1 and MS2 are communicating with two different base stations BS1 and BS2 on adjacent time slots, they still may interfere. This is because MS1 is synchronized with BS1 and since its distance to BS2 is different from its

distance to BS1, the signals from MS1 would have a different delay to BS2 and thus may interfere with the time slot allocated to MS2.



dist(BS1, BS2) = D
dist(MS1, BS1) = d

Figure 1. (*upper*) Even if the BSs are synchronized, the time slots allocated to different BSs may interfere. (*lower*) Distance from MS1 to BS1 is always greater than distance of MS1 to BS2 minus distance of BS1 to BS2.

On the uplink: the duration of interference from MS1 on BS2 can be computed as the time it takes to reach BS2 from MS1 minus the time it takes to reach BS1 from MS1 location. This is because mobiles communicating with BS2 are assumed synchronized to reach BS2 with zero delay. The worst case happens when *distance(MS1,BS2)-distance(MS1,BS1)* is maximal, while MS1 is still at distance less than BS1 cell radius (see Figure 1 (right)). This distance is bounded by *distance(BS1, BS2)*. The bound is reached when MS1 is aligned with BS1 and BS2 on the right of BS1. Thus the maximum duration of interference is $\frac{c}{\text{distance}(BS1, BS2)}$.

On the downlink: this is exactly the symmetric case of the uplink.

B. SOLUTION TO FREQUENCY SHARING

There are two possibilities to avoid interference between cells: by using the guard time between slots, or by using a special partitioning of the frame to reduce the number of interfering slots.

USING GUARD TIME:

When the distance between two BSs is less than a set limit, the guard time between slots may be enough to avoid interference. Given the guard time, we can deduce the limit of distance separating two BSs that can use adjacent time slots. FMA1 proposes to partition a frame in 8 time slots, each 1/8 time slot can be subdivided into 2 slots (1/16). The 1/16 slots can also be subdivided into 4 (1/64) slots. The 1/8 slots are used in CDMA mode with a spreading code of length 16 (while the 1/16 and 1/64 slots are used without spreading). The maximal distance between BSs for 1/8 slots is around 8 kms and for 1/16 and 1/64 slots this limit is 1.2 kms.

Burst	Guard time	Critical path length difference (guard time*c)
1/8 slot 1	26.8 μs	8.04 km
1/8 slot 2	25.4 μs	7.62 km
1/16 slot	4.23 μs	1.27 km
1/64 slot	4.032 μs	1.21 km

Table 1. FRAMES FMA1: Bursts guard and maximum distance between BS for frame sharing.

A first result is that in a given macro-cell C_M the sublayers cells that are located at a distance (from C_M center) less than the critical limit can share the same frequency by partitioning the set of available slots. An additional remark is that a slot used by a micro-cell on the downlink never creates interference on the following slot even if used by a macro-cell. The reason is that the BSs of the micro-cell and macro-cell are either separated by a distance less than 8 km and there is no problem or the distance is greater than 8 km and the interference created by the micro-cell can be neglected when it reaches the macro-cell BS. Using the same reasonning and since BSs are synchronized we induce the following: the macro-cell, when using a slot for the downlink, never creates interference on a preceding slot used by a micro-cell.

FRAME PARTITIONING:

To reduce the number of interfering slots, the frame could be partitioned in a special manner. The micro-cells could be taken starting from the last slot (7) while the macro-cells slots are taken starting from the first slot (0). With this partitioning technique there is only two slots which may experience interference. It is the first slot after the *macro/micro slots border* and the last frame slot.

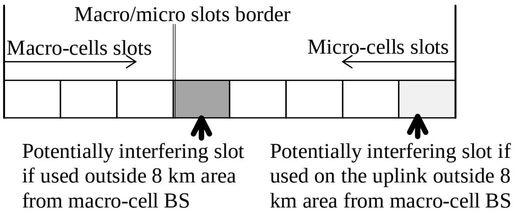


Figure 2. Slots are allocated to macro-cells starting from the first one and to the micro-cells starting from the last one.

COMBINING GUARD TIME AND PARTITIONING:

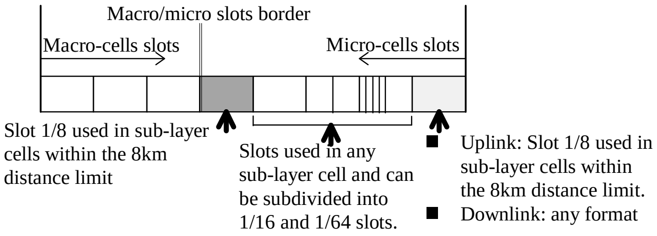


Figure 3. Shows how the frame slots can be allocated to macro/micro cells.

Combining the constraints on guard time and partitioning we can propose a slot allocation mechanism that reduces interference. Slots are allocated to the macro-cell starting from the first one and to micro-cells starting from the last one. The slot following the macro/micro slots border can only be used by micro-cells whose BS is within the 8 km distance limit. This slot cannot be further subdivided into 1/16 or 1/64 slots, unless if one accepts to loose the first sub-slot and use it as a large guard time. The other slots (except the last frame slot) allocated to micro-cells can be subdivided into 1/16 or 1/64 slots if the distance between micro-cell BSs is less than 1.2 km (which is generally true since the limit on micro-cells radius is 1 km and specially for HCS environments where micro-cells are much smaller). The last frame slot can be used on the downlink and subdivided without any constraint or used on the uplink as an 1/8 slot within the 8 km distance limit. With this technique no slot is lost. This is because it is enough for a slot to be used in at least one micro-cell to have the same efficiency as using it by the macro-cell, and since several micro-cells can exist within an 8 kms radius area, such system can be seen as without slot loss.

In [6] it was found that a scheme where layers that are separeted in frequency gives a better performance than a scheme where all slots are available to all the layers. However, this result was due to the fact that at the level of micro-cells a slot in a frequency band shared between the two layers was used with a reuse factor equal R_M*R_m (where R_M is the reuse factor applied to macro-cells and R_m is the reuse factor applied to micro-cells). Implying that the macro-cells have a reuse factor R_M*R_m which is much lower than R_m the value used when the layers are separated in frequency. However, when the BSs are synchronized there is no reason not to reuse a time slot currently used in micro-cells underlying C_M in micro-cells underlying C_M' and which is adjacent to C_M. Thus, a scheme where frequency bands are shared between layers gives better flexibility and improves the capacity of the overall system to handle traffic. This is specially true for FMA1 where the channels are around 2 MHz wide.

III. RESOURCE MANAGEMENT ALGORITHMS

The introduction of the HCS has an impact on several resource management algorithms. It requires these algorithms to take into account a new dimension. Thus, for Handover/Admission control, the algorithms have to select the right layer and the right cell. Congestion/Load control can be optimized to efficiently use the resources (e.g., MS connections are moved to the lowest layers whenever possible). Furthermore, since for non-TDMA systems, the HCS scheme is based on frequency splitting between layers, an interlayer slow DCA algorithm has to be provided to optimize the resources splitting.

HCS resource management algorithms require a speed estimation measurement from the mobile. The speed estimation algorithm can be either based on Doppler-shift (level crossing rate (LCR), zero crossing rate (ZCR) [3, 4]), slope estimation [5] or cell crossing rate. The advantage of the last technique for speed estimation is that it is directly related to the number of handovers executed by the MS, but it may be too slow to adapt to rapid changes of speed. Resource management may be based on both techniques. It can be noticed that the introduction of the HCS concept does not affect the power control algorithm. The same intra-layer power control algorithms are used inside each layer.

A. HANDOVER (HO)

The intra-layer HO algorithms aims at providing the required quality while minimizing the power emission of the MS/BS. The inter-layer HO algorithms have to incorporate some additional parameters that are specific to HCS. The most important one is to minimize the waste of resources. Thus, MSs are handed-over to lower layer cells whenever possible unless if: the MS speed is too high and would require too many HO per unit of time or if the lower-layer cells are congested and cannot service the MS.

HANDOVER ALGORITHM DESCRIPTION:

The handover can be either MS initiated or network initiated. In this section we will focus on MS initiated handover. The network initiated handover will happen when a cell is congested and will be treated in the Admission and Congestion Control section.

There are two types of HCS handover: intra-layer handover and inter-layer handover.

- Intra-layer handover is initiated for the same reason as in a single layer network. It can be based on pathloss, uplink interference or downlink transmission power. The intra-layer HO algorithms are the same as if only one layer existed. When the target cells in the current layer are congested the intra-layer HO leads to an inter-layer HO.

- Inter-layer handover is initiated when the mobile is entering a region that is not covered by cells of the same layer or when the speed of the mobile does not correspond to the layer speed range. An inter-layer handover can also be initiated by the network for congestion and load control reasons.

The mobile measures the pathloss to all candidate base stations and measures its own speed (or number of cells in the current layer crossed every unit of time). The measurements are transmitted to the BS when the pathloss or the speed substantially changes (exceeds the layer threshold or becomes acceptable for a lower layer cell) or if requested by the BS (e.g., when a cell is congested it may want to hand-over some of the services MSs to different cells or layers). Another possibility is to periodically transmit the measurements.

For each layer the mobile maintains an active set of BSs (ActiveSet(L_i)). It represents the set of cells to which the MS can execute a HO. The radio network controller (RNC) makes the decision to which cell the MS will execute a HO. This decision can be based on an algorithm such as:

```
Algorithm Handover()
{
if ( MobileSpeed ∈ LayerSpeedInterval(currentLayer) and
     ActiveSet(MS, currentLayer) ≠ {})
then      if (IntraLayerHO is possible)     /* no cell congestion */
          then ExecuteIntraLayerHO(MS);
          else if (InterLayerHO(MS) is possible)
                then ExecuteInterLayerHO(MS);
                else QueueHORqst (MS);

else-if (InterLayerHO(MS) is possible)
     then ExecuteInterLayerHO(MS);
     else QueueHORqst(MS);
}
end-algorithm
```

Algorithm 1. When the MS speed is still in the layer interval of accepted speed (+/- threshold) an intra-layer HO is first tried. If the intra-layer HO is not possible or the speed is outside the layer interval an inter-layer HO is initiated. Combined with the speed, the average time between two HO could be used.

```
Algorithm SimpleInterLayerHandover(MS)
{
L = LowestAcceptableLayer(Speed(MS));
if ((∃ L'≥L and ∃ C ∈ ActiveSet(MS, L')) such that NonCongested(C))
then return(yes);
else return(no);
}
end-algorithm
```

Algorithm 2. This algorithm finds the closest layer L' to the suitable layer L (MS is suitable to L because of speed and average time between HO) such that L' contains a cell that can service MS. It takes into account that some MSs are not covered by higher layers (e.g., when a MS is in a pico-cell it may be not possible to service it by a micro-cell).

Other more elaborated *InterLayerHandover* algorithms may be developed. For example, a simple HO algorithm may not find a cell that has enough available resources to accept the MS HO request, however by handing-over (inter-layer or intra-layer) some other mobile MS', it may be possible to accept MS. Another possibility is to "sub-rate" other connections (if the HCS is not congested) to free some resources, then the congestion control mechanisms would slowly regulate the whole hierarchy.

The handover requests are queued when they cannot be executed. For an intra-layer handover priority is given to MSs with high speed. For an inter-layer handover, priority is given to MSs that cannot be serviced anymore by its current layer.

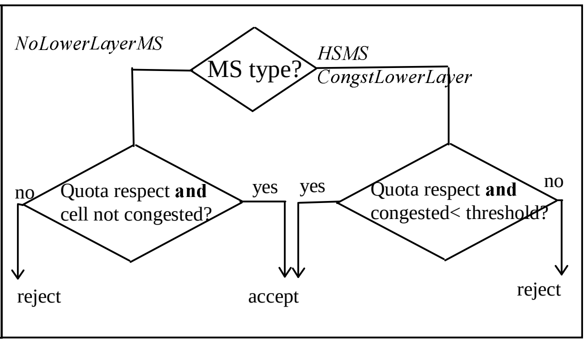
B. ADMISSION/CONGESTION CONTROL (AC/CC))

The HCS congestion control algorithms aims at regulating the traffic between layers. In each layer, the algorithms would give preference first to the mobiles that cannot be serviced by lower layers, then to mobiles for which all the lower cells that may service them are congested ; finally to high speed users that would induce too many handovers if left on the same layer. At each layer, some resources can be reserved for the first type of mobiles and for the two other types of users. The other resources are available for all mobiles. If we want to give priorities to MSs, the admission control accepts only the mobiles that can only be serviced by the cell when it is near to be congested.

ADMISSION CONTROL (AC):

In this section we present an example of AC algorithms. First, an the admission control for Handover, then an admission control for call setup. In the first algorithm, each cell has two control parameters: *PureCellLayerReservedResources(C)*, *HSMSandCLLCReservedResources(C)*. These resources are reserved for specific MSs (e.g., *PureCellLayerReservedResources(C)* = 1/3, *HSMSandCLLCReservedResources(C)* = 1/6). We can differentiate between 3 types of MSs: *NoLowerLayerMS*, *CongestedLowerLayerMS*, *HSMS* (High Speed Users: users that can be serviced by lower layer cells but would induce too many HO). The AC algorithm favors *NoLowerLayerMS* by not rejecting MSs unless if the free resources exceeds a congestion threshold.

```
Algorithm AdmissionControlforHO(MS)
{ if (MS is NoLowerLayerMS)
 then     if ((HSMS and CongestedLowerLayerMS quota is respected)
            and (cell is not congested))
          then Accept;
          else Reject;
else  if ((NoLowerLayerMS quota is respected) and
         (FreeResources > CongestionThreshold$_{HSMS-CongestedLowerLayerMS}$ ))
          then Accept;
          else Reject;}
end-algorithm
```



Algorithm 3. Admission control algorithm for handover.

The admission control for call setup is similar to the one for handover, with the only difference that a call setup is not accepted when a cell is near to be congested. The cells keep few resources for HO.

LOAD/CONGESTION CONTROL:

The load control algorithm continuously optimizes the resources utilization. A MS is handed over to a lower layer cell or to a less congested cell whenever possible. To avoid instabilities waits until the system reaches a stable state after every $N_{MaxSuccessiveILHO}$ inter-layer HO executed by the load control.

```
Algorithm InterLayerLoadControl()
{/* InterLayerHO for resource optimization */
for (i=0; i< $N_{MaxSuccessiveILHO}$; i++)
{
Let S = {MS | ∃L'<L=Layer(MS) and ∃ C∈ ActiveSet(MS, L')
            and RqstResources(MS) < FreeResource(C)};
Let MS ∈ S | L'-L is maximal;
InterLayerHO(MS);
}
IntraLayerLoadControl(cells controlled by the RNC);
Wait (until: system is stable);}
```

Algorithm 4. Load control algorithm.

IV. CONCLUSION

We have shown that resource sharing in an HCS environment is not straightforward but can be achieved under some circumstances to provide efficient HCS implementation. Further, research can be done for improving/adapting such a scheme (e.g., in a time division duplex mode). We also provided several inter-layer algorithms for resource management. Further analysis of the algorithms (e.g., call/handover blocking probability, call blocking, rate of handover) still has to be done.

ACKNOWLEDGEMENT

This work has been performed in the framework of the ACTS AC090 FRAMES, which is partly funded by the European Commission. The author would like to acknowledge the contributions of colleagues from Siemens AG, Roke Manor Research Limited, Ericsson Radio Systems AB, Nokia Corporation, Technical University of Delft, University of Oulu, France Telecom CNET, Centre Suisse d'Electronique et de Microtechniques SA, ETHZ, University Kaiserslautern, Chalmers University of Technology, The Royal Institute of Technology, Instituto Superior Técnico and Integration y Systema.

REFERENCES

[1] E. Nikula et al. "Flexibility in Providing UMTS Services with FRAMES FMA1", In proceedings of ACTS Mobile Communications Summit'97, Oct. 7-10, pp690-696, Aalborg, Denmark.

[2] Robert S. Karlsson and Jens Zander "On Hierarchical Cell Structures for Future Wideband Wireless Access", In proceedings of conference Nordic Radio Seminarium 1997, May 1997.

[3] W. C. Lee, "Mobile Communications Design Fundamentals", 1993, John Wiley & Sons.

[4] K. Kawabata, T. Nakamura, and E. Fukuda, "Estimating Velocity Using Diversity Reception", Proceedings of Vehicular Technology Conference: VTC'94, 1994, Stockholm, pp371-374.

[5] M. Scheibenbogen, M. Schäfer, "Handover in Hierarchical Networks Supported by User Speed Estimation in Microcellular Environment", ITG-Fachbericht Mobile Kommunikation, 1995.

[6] I. Chih Lin, L.J. Greenstein and R.D. Gitlin, "A Microcell/Macrocell Cellular Architecture for Low- and High-Mobility Wireless Users" IEEE Journal on Selected Areas in Communications, Vol 11, No. 6, 1993, pp885-891.