Summary of Resource Management for QoS Support in Cellular/WLAN Interworking

Between 1995 and 2005, wireless access network was improved dramatically which led to them becoming very prevalent. At the time, cellular networks and wireless local area networks (WLANs) showed and had the most potential.[[1]](#endnote-1) Transforming us from analog to digital technology, came the second-generation (2G) cellular networks. Then came the 2.5G cellular networks which provided low-rate packet-switched data services. After the 2.5G cellular networks, third-generation (3G) cellular networks with 2Mb/s as the maximum bit rate came to life because of the unprecedented demand for high-rate multimedia services. Between WLANs and cellular networks, WLANs displayed the ability to provide lower cost and higher rate data services at an 11 Mb/s and 54 Mb/s maximum rate – depending on where you are – while cellular networks had wide area coverage. The plan was to make 4G wireless networks heterogeneous to provide seamless Internet access for mobile users with multimode access capability.

A major challenge with cellular/WLAN interworking is how to take advantage of cellular networks’ wide coverage and roaming support, and WLANs’ high data rates. Other challenges include resource allocation and CAC.

Based on the interdependence between the two access networks, the interworking architectures can be classified into two categories: tight coupling architecture and loose coupling architecture. On one hand, in tight coupling architecture, the WLAN is connected to the cellular core network and appears to the cellular core network as one cellular radio access network. A user roaming across the two domains is based on the mobility management protocols of the cellular networks, thus enhancing the interdomain mobility management capability. However, there are disadvantages to the tight coupling approach: an interface in the cellular core network exposed to WLANs is required, cellular core networks will probably go through a network bottleneck because a large volume of WLAN traffic will go through it, and WLANs and cellular network will need to have compatible protocol stacks.

On the other hand, loose coupling architecture directly connects WLANs to the internet backbone without a direct link between WLANs and cellular core network. Its main advantage is the deployment of both domains. The disadvantage is that mobility signaling may traverse a relatively long path.

to interconnect heterogeneous IP-based wireless access networks with the Internet backbone in 4G networks, it is well recognized that an all-IP DiffServ platform is the most promising architecture to provision broadband seamless global access.

Loosely coupled interworking of cellular networks and WLANs matches well with the emerging evolution toward an all-IP 4G infrastructure and can be naturally implemented in a domain-based DiffServ platform.

The base station (BS) or access point (AP) provides the mobile station (MS) with Internet access. All the wireless domains are interconnected through the DiffServ Internet backbone to provide end-to-end Internet services to an MS.

For cellular networks, resource allocation plays a vital role in effectively provisioning QoS guarantee to each MS and efficiently utilizing scarce radio resources. However, when cellular networks are integrated with WLANs, it becomes much more challenging to achieve QoS provisioning and efficient resource utilization. There are factors that we need to consider if we want to create an effective resource allocation solution.

Resource allocation solutions differ in cellular networks and WLANs. In cellular networks, based on a centralized architecture, the BS has the ability to provide a QoS guarantee to MSs by properly scheduling their access to the wireless channel. schedulers located in different BSs can also coordinate with each other. in the current WLAN, two-channel access functions are defined: the mandatory distributed coordination function (DCF) and optional point coordination function (PCF) in a centrally controlled manner.

both cellular access and WLAN access are available to MSs within WLAN-covered areas. the handoff from a cellular network to its overlaid WLANs is optional, and mainly happens to enhance QoS, lower cost, and balance traffic load. The handoff impacts resource allocation and QoS guarantee. Vertical handoff: handoff between cellular network and WLAN. Horizontal handoff: occurs between BSs and APs within a homogeneous wireless network.

When it comes to dynamic user mobility, the handoff procedure may cause extra delay, packet losses, and even connection interruption. This case would be much worse with the vertical handoff. It is desired to assign MSs with high mobility to a network with large coverage. Users within the areas of low mobility.

As for multiple service types, real-time services are sensitive to delay, while the main concern for delay-tolerant data service is throughput. Taking advantage of a centralized architecture, cellular networks can serve real-time traffic effectively. In contrast, it is difficult for WLAN to meet the strict delay requirements of real-time services. When it comes to serving bursty data traffic, WLANs are more efficient than cellular networks. The mainstream cellular networks cannot handle traffic load asymmetry. WLANs can be viewed as operating in a virtual time-division duplexing (TDD) mode, which can effectively handle the load asymmetry of data service.

Integrating WLAN and cellular networks will produce new challenges when it comes to resource allocation. Various new factors should be taken into account to develop an effective admission scheme for cellular/WLAN interworking. When an incoming service requests into a WLAN-covered area, a decision needs to be made on whether to admit the incoming call to the WLAN or to the overlaying cell. The decision can have a significant impact on resource utilization efficiency and QoS satisfaction. If the total resources of the two tiers are allocated to users by jointly considering factors such as available capacity, traffic characteristics, user mobility, and QoS support capability, higher utilization and better QoS assurance can be achieved. handoff traffic should be differentiated from new traffic in terms of call admission. When a user moves from an area with only cellular coverage to an overlaid WLAN area, the ongoing call of the user can be handed over to the WLAN for balancing load, lowering cost, and so on. Unless there is no space capacity. In that case, the call should remain in the cellular network. On the other hand, for a new call originating within a WLAN-covered area, either the covering cell or the WLAN is first selected for admission according to the service type, user mobility, and current network status. If the service request is rejected by the preferred network, it can overflow to the other network for admission.

When it comes to resource sharing between voice and data services, restricted access mechanism is used to properly share the total bandwidth between voice and data services. Voice traffic is offered preemptive priority over data traffic and occupies up to a certain amount of bandwidth to meet its strict QoS requirements and the remaining bandwidth is dedicated to data traffic. All bandwidth unused by current voice traffic is shared equally by ongoing data flows. to strike a good balance between high utilization and fine QoS guarantee, the number of data calls admitted in a cell or WLAN should be restricted, so that the constraint for mean data transfer time can be satisfied.

voice calls are admitted with a preference to the cellular network To minimize the impact of latency and processing overhead induced by frequent vertical handoff and because it can provide the fine QoS provisioning that is required by voice traffic to meet its strict delay requirement. data traffic has a better rate adaptation capability. The larger bandwidth provided by a WLAN allows for the consumption of the allocated resources for less time. In an area of only cellular coverage, new voice and data calls must request admission to the covering cell. In an area of WLAN coverage “double-coverage area”, there are two choices for the new voice calls: A new voice call will first try to get admission to the cell. The request is rejected if there is not enough bandwidth to accommodate the new voice call. The call will leave the system when both the covering cell and WLAN reject the voice service request. The first admission choice for new data calls in the double-coverage area is the WLAN. If the admission request of a data call is rejected by the WLAN it will not try the cell, as the overflow of data traffic to the cellular part brings little benefit to data calls. To provide priority to voice (handoff and new) calls in the cellular-only area, a limited fractional guard channel (LFG) policy [10] is used. A certain amount of cell bandwidth is reserved for prioritized voice traffic. For an ongoing voice or data call moving out of the WLAN or from one cell to another within the cellular network, a handoff should proceed, otherwise, the call will be dropped. For ongoing calls from a cell to its overlaid WLANs, If the mean transfer time for data calls is not violated when the incoming data call is admitted to the WLAN, the data call is handed over from the cell to the WLAN; otherwise, it will remain in the cell. For an ongoing voice call, no handoff to the WLAN will proceed so that the voice call remains in the cellular network.

To evaluate the performance, we assume that voice and data call arrivals are Poisson. The voice traffic model is relatively simple as the voice call duration is exponential. On the contrary, data calls are more elastic to bandwidth variations, and the data call duration depends on the data file size, allocated bandwidth to the data call, and system steady-state distribution. Measurements have demonstrated that the data file size follows heavy-tailed distributions such as lognormal and Pareto distributions. By properly setting the effective bandwidths of voice calls and data calls, packet-level QoS such as packet delay and packet loss can be guaranteed, as long as the allocated bandwidth to a voice or data call is no less than the corresponding effective bandwidth requirement. Due to the coupling between the cellular network and WLANs in terms of resource allocation, intuitively an increase in the admission region of one network may result in a reduction of that in the other.

User mobility within the double-coverage area can be characterized by the WLAN residence time.

Compared to the WLAN-first scheme. The average handoff number of a voice call is significantly reduced, which means much less latency and fewer packet losses due to handoffs. The improvement increases with higher user mobility.

In conclusion, A new admission strategy for integrated voice and data services has been proposed, which prefers the cellular network for voice service and WLANs for data service, according to the characteristics of the cellular network and WLANs, the distinct features of voice and data traffic and their QoS requirements, and user mobility patterns. The admission strategy can effectively improve the performance of high-priority voice service and, at the same time, fully utilize the large bandwidth of WLANs for data service.

1. Wei Song, Hai Jiang, Weihua Zhuang, and Xuemin (Sherman) Shen. Resource Management for QoS Support in

   Cellular/WLAN Interworking University of Waterloo, 2005, 1 [↑](#endnote-ref-1)