

BEYOND 3G WIRELESS NETWORK DESIGN FOR OPTIMAL RESOURCE UTILIZATION

Ahmed Zahran, Ben Liang

Aladdin Saleh

Department of Electrical and Computer Engineering
University of Toronto
Email: [zahran, liang]@comm.utoronto.ca

Wireless Technology Department
Bell Canada
Email: saleh@bell.ca

ABSTRACT

Efficient resource utilization is a crucial goal for network operators to achieve good return on their investments. In next-generation multi-service wireless network, service integration of different access technologies such as WLAN and 3G cellular networks increases system complexity and makes system design a more tedious task especially with the embedded heterogeneity of access technologies and applications. This paper presents an optimal resource utilization problem for an integrated 3G-WLAN system. The problem solution is used to determine the values of different system parameters to achieve efficient resource utilization without resource over-provisioning. Additionally, we discuss how better resource utilization can be realized by tuning roaming policies using techniques such as dynamic pricing.

1. INTRODUCTION

Next generation wireless communication network is featuring tremendous growth in wireless resource demand due to the huge increase in the number of network subscribers and the introduction of new revenue generating resource demanding services, such as multimedia-based applications. Although network upgrade is always a candidate solution for this problem, operators may find it economically inefficient. Recently, the service integration of different pervasive radio access technologies evolves as an alternative solution that improves both network resource utilization and user perceived quality of service (QoS). However, the great characteristic-diversity of access technologies and services makes system design a real challenge to achieve efficient resource utilization.

The integration of third-generation (3G) cellular networks and wireless local area networks (WLANs) has recently received wide support from standardization bodies and network operators [1, 2]. The former provide universal coverage with limited bandwidth and high cost service, while the latter provide large bandwidth in limited area, known as hotspots, at low cost. Combining these characteristics with the clustered nature of heavy demanding applications in areas such as airports, coffee shops, and other public areas, where WLANs are usually deployed, draws a win-win scenario for all key players. In this scenario, cellular operators conserve precious wireless resources by offloading a huge amount of their traffic to WLANs; while in the mean time, users will enjoy cost reduction while keeping themselves always connected to the best available network [3] in their vicinity.

In the integrated WLANs-3G cellular networks model, users transfer their utilized resources as they roam from one access technology to another, known as vertical handoff (VHO) [4]. This VHO greatly impacts session dynamics and network resource utilization depending on several factors such as available network resources, active user density, user mobility and other factors that are interwoven with each other making system design a real technical challenge. More details about VHO is discussed in Section 2. In this paper, we study efficient network resource utilization as a crucial goal from the network operator's point of view due to its relation to the return on investment. First, we present an optimal resource utilization problem in Section 3. This problem considers both bandwidth and processing power as the main system resources. Numerical results and design guidelines are then presented in Section 4. Finally, we conclude and present the future work in Section 5.

2. BACKGROUND AND RELATED WORK

The integration of WLANs and 3G cellular networks has become recently a subject of great interest. Several system design issues have been considered such as the integrated architecture, mobility management, and security. The architecture choice is a centric issue on which several design choices depend. Loosely-coupled architecture [1] that assumes independent implementation of different networks is receiving a wide support due to its flexibility and its compatibility with the next-generation network's ALL-IP vision. In this architecture, Mobile IP (MIP) [5] is expected to be the main mobility engine due to its maturity and completeness when compared to other mobility management solutions. In its basic version, MIP introduces two new network entities for mobility support known as home agent (HA) and foreign agent. The former is employed for MT registration and traffic forwarding whenever it roams to a foreign network, while the latter advertises for foreign domains and takes part in the registration process.

As the mobile terminal (MT) traverses the network, VHO decisions are taken by a pre-installed mobility management software that follows a predefined handoff policy. However, underlay networks, such as WLANs, can be safely considered the preferred network by system users due to its lower cost and larger bandwidth. Hence, MT are expected to *always* VHO to WLANs whenever encountered. Generally, VHO is accompanied by several operations such as resource allocation and traffic redirection that have great impact on sys-

tem resource utilization especially in a multi-service heterogeneous network, such as the integrated WLAN-3G model. For example, in the MIP paradigm the mainly utilized resource alternates between bandwidth and processing resources, as the MT is handed off between WLAN and the cellular network. This alternation greatly impacts the utilization of different resources. In this context, resource utilization is defined as the percentage of the used network resources to the available resources.

In next-generation heterogeneous multi-service wireless systems, resource utilization depends on several factors including but not limited to application demands, user density, available network resources, and application penetration ratios, defined as the percentage of active sessions of a specific application to the total number of active sessions. All the aforementioned factors are interwoven with each other making efficient system design a real technical challenge that requires accurate mobility and traffic models for proper performance evaluation and system design. Based on our previously developed Coxian-structure-based mobility and session models [6] for next-generation network, we study the impact of VHO on efficient cellular resource utilization considering different implementation scenarios, network resources, and VHO strategies as shown in Section 3. In this work, we utilize some performance metrics that are derived in [6] including cellular utilization time (L_c), WLAN utilization time (L_w) and VHO rate (N_{HO}). The derivations of these parameters are not presented here due to space limitation; however, interested readers are referred to [6] for these details.

3. OPTIMAL RESOURCE UTILIZATION PROBLEM

In our study, we assume a 3G network that provides a universal service with WLANs representing high data rate pockets inside the 3G-network, and we assume that MTs are capable of communicating with both access technologies. WLANs overlaps with P_{ow} of the cellular network, while P_{oc} of the network has a unique cellular coverage, where $P_{oc} + P_{ow} = 1$. Another assumption is that users always VHO to WLAN pockets whenever they are encountered. Additionally, we assume that the network provides the users with a set of m different services. Each application i has bandwidth requirements of b_c^i and b_w^i with corresponding average session holding time of t_c^i and t_w^i for cellular network and WLAN respectively. The values of these parameters vary depending on the application nature as will be shown the following section. In this model, the main system resources are the 3G cell allocated bandwidth and the HA processing power (HAPP) defined in terms of its packet processing capacity. Additionally, we assume that each active user is running only one application i ; and consequently, application i penetration, x^i , is calculated as the ratio of the active users using application i to the total number of active users.

Based on the resource utilization definition, we express cellular bandwidth utilization, U_c , as

$$U_c = \sum_{i=1}^m \frac{\rho P_{oc} x^i b_{ceff}^i}{B}, \quad (1)$$

where ρ represents the cell active user density, B is the avail-

able cell bandwidth, b_{ceff}^i is the user effective cellular bandwidth utilization that equals $\frac{b_c^i L_c^i}{L_s^i}$ in which L_c^i represents application i cellular utilization time within a 3G cell and L_s^i is application i total session holding time within a 3G cell. In this expression, $\rho P_{oc} x^i$ is the number of users using application i in the cellular part. Consequently, $\rho P_{oc} x^i b_{ceff}^i$ is the amount of wireless resources allocated for application i users. In addition, we express HA processing resource utilization, U_{HA} , as

$$U_{HA} = \sum_{i=1}^m \frac{\rho P_{ow}}{HAPP} \left[\frac{b_{weff}^i}{s^i} c_f + \frac{N_{HO}^i}{L_s^i} c_m \right] x^i,$$

where b_{weff}^i is the user effective WLAN bandwidth utilization and equals $\frac{b_w^i L_w^i}{L_s^i}$ in which L_w^i is application i WLAN utilization time within a 3G cell, s^i is application i average packet size, c_f is packet forwarding processing cost, N_{HO}^i represents application i VHO rate within a 3G cell, and c_m is mobility registration/deregistration processing cost. The first part of U_{HA} expression represents processing resources used for packet forwarding to application i in a WLAN, while the second term represents the processing resources used for handling VHO registration/deregistration requests of MIP.

Hence, we define the optimal resource utilization problem as

$$\max_{\mathbf{x}} (A_{HA} U_{HA} + A_c U_c) \quad (2)$$

such that

$$\sum_{i=1}^m x^i = 1, \quad (3)$$

$$U_c \leq U_{cmax}, \quad (4)$$

$$U_{HA} \leq U_{HAmax}, \quad (5)$$

$$\mathbf{x}_{lb} \leq \mathbf{x} \leq \mathbf{x}_{ub}. \quad (6)$$

where A_{HA} and A_c represent utilization coefficients used to adjust the priority of a specific resource. Here, U_{cmax} represents the upper bound for wireless resource utilization for the specified set of applications. This parameter may be set to a value less than one to avoid the resource starvation of other services. In addition, U_{HAmax} represents the upper bound for HA resource utilization; it is highly recommended that this parameter is less than one to avoid long delays that result from comparable arrival and service rates. \mathbf{x}_{lb}^i and \mathbf{x}_{ub}^i represent a lower bound and upper bound vectors for the service which may be set by the operator to ensure the presence of specific services using the lower bound or to limit the penetration of specific application using the upper bound.

Clearly, the objective function and the constraints are linear functions in x^i ; consequently, this problem is a linear programming optimization problem [7]. There is no simple analytical formula for the solution of a linear program, but there are a variety of very effective methods for solving the problem. In this work, the problem is solved using the MATLAB *linprog* function, which is based on an interior point technique that is well suited to efficiently obtain solutions for linear programming problems.

Table 1. Application Parameters

Unit Parameter	VoIP	V-conf	RoD	VoD
b_c^i (kbps)	12	64	16	128
b_w^i (kbps)	12	64	90	720
t_c^i (minutes)	3	30	60	90
t_w^i (minutes)	3	30	10	15
s^i	2.4	3.2	2.5	6

4. SYSTEM AND APPLICATION SPECIFIC RESOURCE UTILIZATION

In this study, we consider different multi-media based applications such as voice over IP (VoIP), video conference (V-conf), radio on demand (RoD), video on demand (VoD). VoIP and V-conf are considered conversational applications that are characterized by the same session duration and bandwidth requirement independent of the used network. On the other hand, streaming applications such as RoD and VoD are bandwidth greedy unidirectional applications; hence, they benefit from resource rich networks such as WLANs by increasing the download rate for smoother playback. The aforementioned application parameters are shown in table 1 [8]. In this table, the WLAN bandwidth requirement for greedy applications is assumed to be user profile limited as the WLAN resources is always assumed to be more than its demand since the same area can be serviced by a different number of access points.

We study the impact of several factors including user density, available HAPP, and available wireless resources on network design using the presented optimal resource utilization problem. In the optimization problem, we have set U_{HAMax} to 0.95, U_{cmMax} to one, and all the application penetration rate lower and upper bounds are set to zeros and ones respectively. Additionally, P_{ow} is set to 0.3 with WLANs randomly located within 3G cells, and both c_f and c_m are set to one. The solution of the optimization problem is obtained in two steps

1. The application parameters are passed to the session analytical model to obtain network utilization times (L_c^i and L_w^i) and VHO rate (N_{HO}) for different applications assuming the mobility parameters to be $\alpha = 0.9$, $\mu_v = 1$, and $\sigma_v = 0.5$ [6].
2. These parameters are used to calculate the coefficients of penetration rates of the resource utilization optimization problem that is solved using MATLAB. These results are then used to calculate the HA processing resource and wireless resource utilizations.

Additionally, we assume that both processing and cellular resource utilizations have the same importance, Hence, we set both A_C and A_{HA} to the same value, e.g. $A_C = A_{HA} = 1$.

In the following, we investigate the user density, the wireless resource variation and the HA processing resource impact on the optimal resource utilization. It is worth mentioning that the first two factors are technology dependent and may be limited to specific values according to different standards. For example, in CDMA2000 systems, the cell allocated bandwidth is determined as multiple of 1.25MHz for

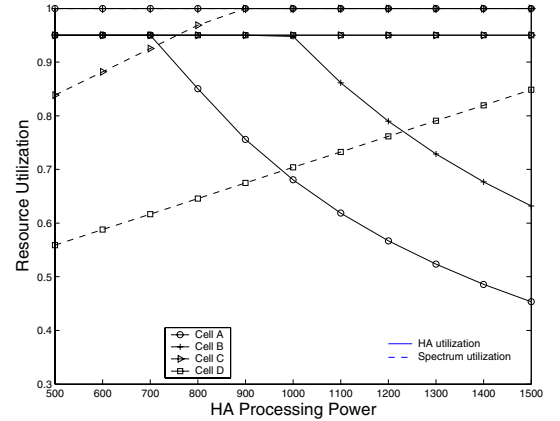


Fig. 1. Resource utilization vs. HAPP for different system capacities

different phases; however, even for the same allocated bandwidth, the service rate may vary depending on the technologies adopted in the system such as adaptive rate modulation, MIMO, and hybrid ARQ [9]. Similarly, maximum user density is a technology dependent factor, e.g. the maximum number of supported users per cell, known as *soft capacity*, depends on the number of simultaneously supported codes in CDMA systems. However, the cell's soft capacity can be altered by using quasi-orthogonal codes at the expense of degraded service quality. Hence, we do not assume a specific system configuration in our study. We consider four different cells, A, B, C, and D, that are allocated different bandwidths of 2Mbps, 4Mbps, 8Mbps, and 12Mbps respectively.

4.1. HAPP Design Values

Figure 1 plots the bandwidth and HA processing resource utilization versus different HAPP as a cell design parameter for different cells. In this figure, the soft capacity is set to 200 users for the four cells. Clearly the figure shows that the bandwidth allocated resources for cells A and B are fully utilized by their users, while the HA resources are fully utilized until specific values for different system, e.g. 700 packets/sec for Cell C, beyond which increasing the HAPP is useless¹. Hence, this figure provide system designers with an upper bound for the required HAPP for specific cells. Additionally, as shown for Cells C and D, the allocated bandwidth resources are under utilized below some HAPP values, e.g. 800 packets/sec for cell C. Hence, this value determines the lower HAPP bound required for full bandwidth utilization under the "always VHO" policy. It is worth mentioning that this value will differ for other VHO policies. To sum up, this figure provides a road map for determining the required HAPP values to enable full resource utilization in heterogeneous systems.

4.2. Bandwidth Utilization and User Density

Figure 2 plots the bandwidth and the HA processing resource

¹For the same set of parameters, full spectrum utilization is achievable for the four cells in the multi-service homogeneous systems.

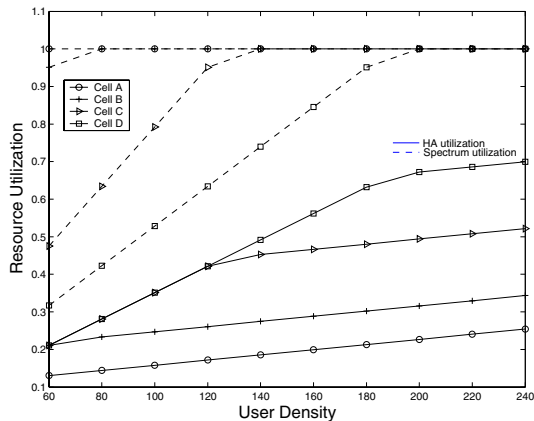


Fig. 2. Resource utilizations vs. user density for different cells

utilizations versus different user densities for different cells. In this figure, we set the HAPP to 3000 pkts/sec as an infinite packet processing resource noting that proper design values can be determined as shown in the previous subsection. Clearly, as shown in the figure, the HA utilization is underutilized for all four cells. Furthermore, the bandwidth utilization of cells B, C, and D is not fully utilized until a specific number of users is reached, e.g. 140 for Cell C. This number can be considered the user density threshold required to ensure full bandwidth utilization under maximum expected load for the “always VHO” policy. Hence, this threshold value should be considered as a main parameter in the cell bandwidth evolutionary upgrade decision. Additionally, these plots can be considered, from another perspective, a chart of the maximum expected load for different user densities.

4.3. Discussion

In addition to the design guidelines that the above problem solution provides, another important outcome of this work is a means to compute the optimal penetration ratios for different multimedia applications. These ratios represent the optimal application penetration distribution that realize the maximum resource utilization. Practically, the real application penetration ratios depend on user demand patterns that may not match the obtained optimal penetration values. Consequently, systems resources may be underutilized. However, dynamic pricing mechanisms are well known as useful tools to control this demand. For example, when the bandwidth resources are underutilized, the cellular network operator may choose to decrease the price of conversational applications; hence, users holding this type of applications will continue using the cellular network even inside WLANs, which have consequently lost the unique cost advantage for this type of applications. Additionally, users may tune their devices to start downloading specific information when the corresponding service price drops below certain predetermined value. Hence, the obtained optimal application penetration ratios with dynamic pricing schemes can work in harmony to enable the network operators to realize the maximum return on investment from their resources. Last but not least, it is worth mentioning that the

presented optimization problem can be easily tuned to consider different operating scenarios and operator choices.

5. CONCLUSIONS AND FUTURE WORK

The service integration of different wireless access technologies is envisioned as a feasible solution to accommodate the increasing demands on wireless resources. In this model, efficient resource utilization requires careful planning of different system design parameters. Using the presented optimal resource utilization problem, the HA processing power bounds are determined in order to achieve efficient resource utilization and avoid resource over provisioning. Additionally, the minimum cell soft capacity required to guarantee efficient resource utilization after system bandwidth upgrade is determined. As a future work, we will consider other mobility management approaches such as the end-to-end approach. Additionally, we will study how the obtained application penetration ratios can be embedded into dynamic pricing schemes to achieve the optimal resource utilization.

6. REFERENCES

- [1] M. M. Buddhikot, G. Chandranmenon, S. Han, Y. W. Lee, S. Miller, and L. Salgarelli, “Integration of 802.11 and third generation wireless data networks,” in *Proc. of IEEE INFOCOM*, San Francisco, US, Apr. 2003, pp. 503 – 512.
- [2] ETSI, “Requirements and architectures for interworking between HIPERLAN/3 and 3rd generation cellular systems,” ETSI TR 101 957, Tech. Rep., Aug. 2001.
- [3] E. Gustafsson and A. Jonsson, “Always best connected,” *IEEE Wireless Communications*, vol. 10, no. 1, pp. 49 – 55, Feb. 2003.
- [4] M. Stemm and R. H. Katz, “Vertical handoffs in wireless overlay networks,” *ACM Mobile Networks and Applications*, vol. 3, no. 4, pp. 335 – 350, 1998.
- [5] C. Perkins and D. Johnson, “Mobility support in IPv6,” Internet Draft, IETF, Mar. 2002.
- [6] A. H. Zahran, B. Liang, and A. Saleh, “Modeling and performance analysis for beyond 3G integrated wireless networks,” in *Proc. IEEE International Conference on Communications (ICC)*, June 2006, [to appear].
- [7] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge university press, 2004.
- [8] F. Velez and L. Correia, “Mobile broadband services: classification, characterization, and deployment scenarios,” *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 142 – 150, Apr. 2002.
- [9] M. Etoh, Ed., *Next Generation Mobile Systems: 3G and Beyond*, 1st ed. John Wiley and Sons, ltd, 2005.