Moving Beyond 4G Wireless Systems

Jim Martin, Rahul Amin
School of Computing
Clemson University
Clemson, SC USA
Jim.martin@cs.clemson.edu

Ahmed Eltawil, Amr Hussien
Department of Electrical and Computer Engineering
University of California
Irvine, CA USA
aeltawil@uci.edu

Abstract— As broadband data further blends with cellular voice, mobile devices will become the dominant portals to the connected world. However current design practices still involve building independent networks that each make their own resource decisions. To move beyond 4G, wireless networks will need to cooperate to provide better performing and more robust wireless services. In this paper, we present a ‘clean-slate’ approach to wireless networking. The approach requires the following assumptions to hold: 1)Incentives are in place motivating independent, autonomous wireless systems (AWSs) to cooperate with each other to provide users with universal broadband wireless coverage; 2)In any given geographic area, a handset might have access to many independent, autonomous wireless networks; 3)Mobile nodes will support a small number of adaptive radios that are capable of operating in a number of supported communication modes and over a range of spectrum. With these liberating assumptions, we present an architecture for a heterogeneous wireless system that moves beyond 4G systems. We present preliminary results from a MATLAB-based simulation designed to demonstrate the essence of the ideas. We identify the key research challenges that must be addressed to move the idea from a concept to a viable network system.

Keywords-heterogeneous wireless networks; reconfigurable radios; 4G wireless systems; cyberinfrastructure;

I. INTRODUCTION

While advances at both the network layer and the physical layer are prompting unprecedented mobile broadband growth, the means by which mobile terminals are offered services have created inefficiencies in both spectrum usage as well as mobile physical resources. The premise behind this issue is that while emerging devices support a multitude of wireless access methods, the current access methods require the user to select the active access network either by purchasing an appropriate handset (and service) or, in the case of multimodal smartphones, by manually selecting the access network. While this approach has worked fairly well for current 2G cellular networks, upcoming 3G and 4G systems will face significant challenges as wireless operators are held accountable for poor performance or inadequate coverage. It is well understood that individual systems that manage blocks of spectrum independently are inevitably operating at suboptimal performance. However, there are precipitous hurdles that the wireless community faces to move forward.

Until recently, technology was the primary impediment to achieving universal, broadband wireless services that involve multiple radio access technologies. Today, the most significant impediments are the after-effects of antiquated government spectrum allocation policies and the resulting economic forces that drive the wireless industry. The effect is that in many geographic areas licensed spectrum is likely to be underutilized [1]. To move beyond current thinking, spectrum, economic, and technical issues must be addressed jointly. This reality has spurred academic interest across several relevant areas including dynamic spectrum management, cognitive radios, and heterogeneous networks spanning both the physical and the network layers.

The physical layer and MAC layers (i.e., the radio) of a wireless node attempts to achieve the best performance within its own network, generally ignoring impacts of co-located wireless networks. This ‘selfish’ behavior will usually not lead to optimal resource usage. Techniques or paradigms such as cooperative communications, cognitive networking, and dynamic spectrum access attempt to improve spectral efficiency through cooperation at the radio level [2]-[20]. At the network level, architectures to support hybrid or heterogeneous networks have been suggested. Recent proposals have been based on the IEEE 802.21 standard which provides a framework to support vertical handoffs transferring a mobile user between two networks that are based on different radio access technologies[23]-[29].

Although emerging 4G networks embody many of these recent advances, current design practices still involve building independent networks. The wireless industry is at a crossroads. To meet the basic requirement for universal coverage, the number of cellular providers is growing smaller while the size of their networks is growing larger. However, users are requiring more bandwidth intensive services that might not exist in all coverage areas. On the other hand, 802.11 networks have proliferated to the point that they are now considered to be a part of the computing (cyber) infrastructure. To augment wireless connectivity, some organizations or agencies have deployed site-wide, campus-wide, or city-wide broadband wireless coverage. These trends and the subsequent impacts on end users are becoming commonplace. Using Clemson University as an example, depending on the specific location on campus, one can find multiple 802.11 networks (some operated by the University, some operated by local public safety, others operated by the adjoining city of Clemson or nearby businesses whose coverage spills out over areas of the campus), coverage from an experimental campus WiMAX network, and multiple 2G/3G networks from cellular providers including AT&T, Sprint, T-Mobile, and Verizon. Over the next
year, the spectrum on campus will become further cluttered as cellular providers deploy 4G networks. While wireless technology is advancing rapidly, research that facilitate cooperation across independent networks is relatively untouched.

In this paper we present an architecture for a wireless system that in our opinion moves beyond the current definition of 4G. Our ideas are to some degree a ‘clean-slate’ idea as the following assumptions must hold:

1. Incentives are in place motivating independent, autonomous wireless systems (AWSs) to cooperate with each other to provide users universal broadband wireless coverage.
2. In any given geographic area, a handset might have access to a large number of AWSs. The use of concurrent, multiple radio links will be standard procedure.
3. Due to cost and battery constraints, mobile nodes will be limited to a small number of adaptive radios that must be capable of operating in a number of supported communication modes and over a range of spectrum.

A. Proposed Example Architecture

Figure 1 illustrates a motivating example for the ideas presented in this paper. Specific types of users, characterized by application domains, observe a unified wireless network, where a mobile node has access to and supports multiple access technologies. The optimal selection of the current active mode (or modes) is performed at the network level. Independent wireless networks (such as commercial cellular providers) cooperate to provide a higher capacity, more robust network than could be provided if each network operated independently. The network implements services based on a hybrid approach to resource allocation. A centralized algorithm manages bandwidth and user resources at a global level. Localized algorithms based on standards based MAC protocols manage bandwidth over small time scales. Allocation decisions will be based on several metrics including a) the QoS requested by the user, b) the current network resources that the user is consuming, c) the anticipated mobility of the user and d) the available networks that a specific user’s hardware supports.

We envision two economic models that could support our idea: a carrier-centric model and an Internet model. In the carrier-centric model, a service provider offers services for specific markets (e.g., commercial wireless access, solutions for public safety). The carrier might own and operate portions of the physical network and possibly broker ‘peering’ arrangements with other wireless providers. Customers subscribe to a single carrier and gain access to any/all resources/services the subscriber has purchased. Voice, video, Internet connectivity, and VPN are the major services that are expected to be provided. An alternative economic model follows the current Internet model: organizations own and operate autonomous networks. Unification occurs through an overlay network that can be achieved through a combination of standard protocols, standard services, and incentive/reward mechanisms that promote peering and collaboration.

The key differences between the carrier-centric and Internet models include how vertical handoffs are initiated and the granularity of resource allocation. In the former, the carrier initiates vertical handoffs that might reflect load balancing or cost optimization. In the Internet model, handoffs might be based on an extended version of mobile IP that supports vertical handoff protocols based on the 802.21 framework. In the carrier model, network resources will be tightly controlled based on centralized resource allocation decisions that achieve the carrier’s policies. In the Internet model, an end-to-end mechanism allows users to compete fairly with other users for available bandwidth.

![System Concept](image)

Figure 1. System Concept

In the remainder of this paper we present a wireless system that embodies these requirements. We limit the scope of this paper to the carrier model. One possible system is described. Further perspectives of the ideas are provided through simulation. Section II presents and relates the relevant background and provides motivations for the work. We provide preliminary results from a simple MATLAB-based simulation in Section III. We then describe the key technical challenges that must be addressed to pave the way for future ‘beyond 4G’ wireless networks in Section IV. We conclude by describing our current project status and possible future directions in Section V.

II. BACKGROUND AND MOTIVATIONS

Significant advances have been made in the last decade in radio technology. Recent work in cross-layer optimization has shown that breaking the rigid OSI layered stack model can enhance both spectral efficiency and application performance [2],[3]. Increased efficiency can be obtained through the use of transmit diversity methods such as multiple input, multiple output (MIMO) techniques [4],[5]. It has been shown that further improvements are obtained when wireless nodes cooperate using techniques such as relays and other forms of MAC/PHY layer cooperative communications [6]-[8]. Dynamic spectrum access (DSA) is another form of cooperation that is likely to be a foundation of future wireless systems [9]-[11]. DSA requires devices to be frequency agile giving rise to software defined radios and cognitive radios as possible implementations.
Software defined radios (SDRs) have been under development for many years, however, they are limited to relatively narrowband systems due to multiple issues including power consumption and available computing cycles [12]-[14]. The challenge in handset design is managing the tradeoff between flexibility in how spectrum is used and in space/power requirements of the platform.

Cognitive radio (CR) represents an approach to radio networks whereas a radio node “can change its transmitter parameters based on interaction with the environment in which it operates” [15]-[18]. An extension to CR is cognitive networking which is defined as “a network with a cognitive process that can perceive current network conditions, and then plan, decide, and act on those conditions” [19]-[21]. One step further, symbiotic networking observes that current wireless networks tries to achieve the best performance within their own network and generally neglects the impact on co-located wireless networks [22]. In other words, radio behaviors are generally selfish and based only on information observed locally. Symbiotic network effectively takes the idea of cross-layer design to the network level – and proposes cooperation across all layers and network boundaries.

While SDRs, CRs, cognitive networking and symbiotic networking focus on improving efficiency from the bottom up, heterogeneous wireless networks represent methods for cooperation driven from the top down. The IEEE 802.21 standard provides a framework to support seamless mobility through networks based on different radio access technologies (RATs) without the need to restart the radio connection every time the mobile moves to a new network [23]. Another relevant standard, IEEE P1900.4, defines building blocks for enabling coordinated network-device distributed decision making which will aid in the optimization of radio resource usage, including spectrum access control, in heterogeneous wireless access networks [24].

A number of architectures to handle heterogeneous wireless networks have been proposed in the literature. Hierarchical resource managers have been proposed by the Common Radio Resource Management, Joint Radio Resource Management and Multi-access Radio Resource Management schemes studied by the 3GPP group [25]-[29]. In these hierarchical schemes, the local resource managers of different wireless technologies interact with a centralized entity to jointly optimize the process of resource allocation. These schemes are similar to our work that proposes central management of resources through an entity such as Global Resource Controller (GRC). The work in [28] is the only work that has specifically proposed the idea of a global resource controller issuing reconfiguration commands to radios. The open research issue is how to formulate a tractable resource allocation problem when subject to an enormous parameter space. Our work is at the intersection of evolving 4G wireless technologies and cooperative networking. The ideas that we present in this paper extends existing research for at least the following reasons:

1. We assume a wide range of wireless networks and infrastructure will exist and that the most flexible, unified system can only be derived when the devices are multimodal and highly reconfigurable.

2. We assume network-wide resource allocation is required. For simplicity, our initial path assumes a single carrier model, however the ultimate vision is a wireless Internet model where devices at any given time might have many connectivity choices and a set of distributed services exist to allow the device to use the available resources in a manner that achieves some understood and agreed upon compromise between fairness and system resources.

III. SYSTEM DESCRIPTION

A. System Model

Figure 2 illustrates the system model. The system consists of devices (also referred to as nodes) that have connectivity to one or more AWSs. Each AWS will have a controller (referred to as an AWSC) that represents all nodes in the AWS and that serves as a gateway connecting the AWS with other AWSs or external networks. A global resource controller (GRC) manages resources in a manner that supplements decisions made by local AWSCs.

Users are presented with a unified network. Depending on node capabilities, users can operate over more than one AWS at any given time. A node’s TCP/IP stack sees a single IP link. The wireless virtual link layer (WVLL) handles packet scheduling over one or more radio links. Packet resequencing, error recovery using ARQ and/or FEC can optionally be implemented over the tunnel. The radio link block pictured in Figure 2 represents the MAC and physical layer that operates over a portion of the spectrum. A radio would be implemented using a combination of custom hardware along with programmable hardware based on technologies such as field programmable gate arrays (FPGAs) or digital signal processors (DSPs). User data is tunneled over the unified network cloud. The GRC (or another located in the backend network) represents the termination point for the tunnel. Nodes must maintain periodic contact with the GRC by sending status update information periodically. The GRC sends network level status and/or resource management control information periodically to the AWSC (and possibly to individual devices). The unified network provides a best effort datagram service specified. Users purchase a service plan that would specify a downstream and upstream service rate. The network could define more complex data services such as a differentiated service offering. We will explore this in a later stage of the research.

B. Radio Capabilities

Radios are either static or capable of reconfiguration. Static radios are equipped with one or more non-reconfigurable radios. A non-reconfigurable radio supports a limited level of adaptive capability, whatever is required by the AWS technology. The position of mobile devices will be tracked by the GRC as location-based management is required. An AWS will likely consist of multiple wireless networks, generally of the same radio access technology. For example, a particular 3G network or a campus-wide 802.11 network would be considered an AWS. The GRC is necessary to facilitate cooperation across AWSs. For example, a node that concurrently supports 802.11 and 3G might be told by the GRC...
to use the 3G link as the control channel and the 802.11 network as the data channel. Alternatively, if the same node moves to a location that has 802.16e coverage, the GRC might initiate a ‘reconfiguration handoff’. For example, as the node moves out of coverage of the 802.11 network and the GRC determines that the node is in coverage of the 802.16e network, the network could issue a reconfiguration command to the node instructing it to reconfigure the radio to 802.16e. A reconfiguration handoff is a vertical handoff that requires a radio to reconfigure itself.

We present simulation results from a scenario involving 10 AWSs. Each is assigned to one of the following radio access technologies: GPRS, EDGE, CDMA 2000, EVDO Rev 0, WCDMA, HSPA, IEEE 802.11 g, IEEE 802.11n, IEEE 802.16e and LTE. For the cellular based AWSs, the base stations are located in the center of the grid. The IEEE 802.11 b/g and IEEE 802.11n APs remain at the edge of the grid as illustrated in Figure 3.

The simulation involves 100 mobile nodes that are initially positioned randomly within the grid. Nodes move using a random waypoint mobility model moving at a moderate velocity of 10 m/s (36 km/hr). Due to this movement, link performance varies over time depending on the node’s location in the grid and scheduling decisions. To strike a balance between update frequency and overhead, the mobile node sends its connectivity parameters for each access technology to the Global Resource Controller every one second. The Global Resource Controller implements the scheduler that allocates each node’s radio the most efficient access technology. To ensure proportional fairness, the scheduler follows a two step approach. In the first step, the scheduler allocates a minimum required throughput of 500 kbps to each node using its best radios. In the second step, the scheduler distributes unused access technology resources to a window of 10 mobile nodes with best connectivity parameters in increments of 100 kbps. The simulation is run for 10,000 seconds and the distribution of throughput per user (bits/s) is compared for three different scenarios. Each scenario is explained next, followed by the results and discussion.

**Scenario 1:** In the first scenario, each node consists of only one non-reconfigurable radio. Each node has one of three access radios: (i) LTE, (ii) IEEE 802.16e, and (iii) IEEE 802.11n. The ratio of nodes having each radio is equally divided into a factor of 1/3. In a 100 node simulation, 33 nodes have the LTE radio, 33 nodes have the IEEE 802.16e radio, and 34 nodes have the IEEE 802.11n radio. This scenario depicts today’s devices that support only a single mode of operation. It acts as the base case that provides comparison against the other two scenarios, and is expected to exhibit the most inefficient use of spectrum.

**Scenario 2:** In the second scenario, each node consists of three non-reconfigurable radios. Each node has three access radios: (i) LTE, (ii) IEEE 802.16e, and (iii) IEEE 802.11n. In a 100 node simulation, there are a total of 300 radios, a third of which can access LTE, IEEE 802.16e and IEEE 802.11n respectively. This scenario depicts the multi-modal trend towards which the

---

Figure 2 System Model

Most research in cognitive radio and cooperative communications is based on localized decisions. While some wireless technologies do support centralized load balancing, much of the relevant research, especially work in the area of handoffs in heterogeneous networks, assumes the ‘trigger’ that initiates a handoff is determined by the node. Our approach is based on a hybrid approach that supports localized decisions but supplemented as needed by the GRC. This approach addresses the following limitations of localized resource management:

- A decision made on local knowledge (e.g., signal strength) might not lead to the best resource allocation decision as the network might be congested or there might be unseen interference issues.
- Local allocation decisions might not reflect current policy the service provider would like to impose.
- A node might not be able to detect all possible AWSs that are available.

**C. Simulation Description**

We developed a MATLAB-based simulation model with sufficient fidelity to demonstrate the possible benefits of the proposed wireless system. The simulation topology consists of a 2000 m * 2000 m grid. Any number of wireless access technologies can be used for connectivity within the grid. The AWS model is parameterized by transmit power, a propagation model, and a mapping between effective receive power at a given location and effective data rate. The latter capability
mobile device industry is approaching. Each multi-modal device can use one radio at any given time, or it can use all three radios at any given time. The scheduler determines which radios are active for the given time frame and how much traffic is supported on each radio.

**Scenario 3:** In the third scenario, each node consists of three radios that can reconfigure themselves to use any of the 10 technologies that are accessible in a given area. In a 100 node simulation, there will be a total of 300 radios which will use any combination of 10 available access technologies. Again, the scheduler determines which radios are active for the given time frame and how much traffic is supported on each radio.

![Figure 3: Coverage of an example simulation topology](image)

**D. Results and Discussion**

The cumulative distribution function of the average throughput each user experiences in the three scenarios described in the simulation description section is presented in Figure 4. The results are dependent on node mobility, access technology coverage range, allocation resources of each access technology and scheduler implementation. The scheduler implemented in this study attempts to maximize network efficiency and maintain fairness across users. Other forms of schedulers might result in different results. We will explore in detail the scheduling problem in future work. For the purposes of this study, our objective was a simple allocation strategy that is sufficient to illustrate the potential benefits of our ideas.

In the first scenario, the average throughput distribution of each user reflects the throughput supported by each individual access technology. Since a node is limited to one access technology, it is allocated resources only when it gets in the coverage range of the technology. A third of the nodes that have the IEEE 802.11n radio get into the coverage range of the IEEE 802.11n AP only a fraction of the simulation time. Hence, a third of the nodes experience average throughput that is under the minimum required throughput of 500 kbps. The second third of the nodes that have the IEEE 802.16e radio are always in coverage range of the IEEE 802.16e base-station. But the resources of IEEE 802.16e are not sufficient to satisfy anything more than 500 kbps per user. So the second third of the nodes have an average throughput of around 500 kbps. The final third of the nodes consist of an LTE radio that is always in the coverage range of the LTE base-station. The LTE base-station has resources to support more than the minimum required 500 kbps per user. The additional resources (above 500 kbps) are distributed to the users that experience the best connecting conditions. Thus, the final third of the nodes have an average throughput in the range of 1.5-3 Mbps based on the conditions each node experiences through the duration of the simulation.

In the second scenario, the average throughput per user is distributed in a much narrower range of 0.6-1.5 Mbps. Each user uses its best radio to support 500 kbps connection. Once each user is allocated 500 kbps, there are always additional resources that the combination of IEEE 802.11n, IEEE 802.16e and LTE possess. These additional resources of each technology are assigned to the users that experience the best connecting conditions, and on average each user gets anywhere from 0.1-1.0 Mbps of these resources before they are exhausted.

![Figure 4. Throughput comparison based on network devices](image)

**IV. TECHNICAL CHALLENGES**

There are clearly significant technical challenges that must be overcome in order to move beyond 4G systems. The key questions include:

- What are the tradeoffs between: power, degree and frequency of reconfigurability, channel depth, and spectrum efficiency?
- What are feasible optimization strategies?
- When many connectivity options are possible to users and service flows, how should resources be allocated?
o Is there an acceptable compromise between fairness and system efficiency?
• What is ideal SDR platform in terms of power consumption, programming model and scalability?
  o Can partial reconfiguration of the hardware architecture lead to a minimization of the active parts running on the hardware at a certain instant in time, thus leading to feasible SDR systems?
  o Can partial reconfiguration minimize the impact to applications caused by reconfiguration handoffs?
• Is there a system architecture that is flexible enough to support a carrier model and also an Internet model?

V. CONCLUSIONS AND FUTURE WORK

In spite of recent advances in wireless technology, wireless networks continue to be designed as independent networks that make resource decisions without considering co-located networks. Cognitive radio, and the more recent idea of cognitive networks address the layer from the perspective of the physical layer (and up the stack). The hierarchical architecture that we have described adopts a top down perspective. Moving beyond 4G networks requires cooperation at the network layers. Preliminary simulation results confirm that significant increase in spectral efficiency can be achieved. In future work, we plan on addressing the key technical issues. Our next step is to explore strategies for optimal resource allocation that take into account rapidly developing advances in reconfigurable radio capabilities.

REFERENCES