

RATIONALE FOR A CLEAN SLATE RADIO ARCHITECTURE

Ronan Farrell, Rudi Villing
National University of Ireland
Maynooth, Ireland

Alexander M. Wyglinski
Worcester Polytechnic Institute
Worcester, MA, USA

Christopher R. Anderson
United States Naval Academy
Annapolis, MD, USA

Jeffrey H. Reed, Philip Balister, Chris Phelps, Tom Tsou, Joseph Gaeddert, Carlos Aguayo
Virginia Polytechnic Institute and State University
Blacksburg, VA, USA

Sven Bilen
Pennsylvania State University
State College, PA, USA

George Nychis
Carnegie Mellon University
Pittsburgh, PA, USA

John Chapin
Vanu Inc.
Cambridge, MA, USA

Behrouz Farhang-Boroujeny, Neil Patwari
University of Utah
Salt Lake City, UT, USA

Jean-Christophe Schiel, François MONTAIGNE
EADS France S.A.S.
Paris, France

Correspondence Emails: {ronan.farrell@eeng.nuim.ie, alexw@ece.wpi.edu}

ABSTRACT

Radio technology has advanced tremendously over the past century with increased capacities and addressing a disparate range of applications. However, from an external perspective, the design philosophies behind radio architectures and wireless standards remain rooted in the methodologies of the 1930's. In this paper, we propose that this conservative approach has served the communication system engineer well for the past century, but it is now constraining the development of future communication devices and networks that are based on software defined radio technology. The scope for evolutionary improvement on the existing architectures is becoming limited, and it is now appropriate to reconsider our basic assumptions in order to determine whether a radically different approach may yield significant benefits in terms of system performance, cost, mobility, and functionality.

1. INTRODUCTION

Since the first uses of radio in the 1890's, there has been a proliferation of radio applications and an increasing

sophistication of techniques for enhancing transmission of information. An example of this sophistication is the increasing adoption of software defined radio technologies and the related area of cognitive radio systems. Software defined radio, and related technologies, offers the potential for creating fully customizable radio systems which are capable of adjusting to different services as needed by the network or the user, whether those services are at a different frequency or employ a different waveform. Since the concept was first proposed by Mitola in 1991 [1], significant advances have been achieved. Nevertheless, practitioners will readily attest that developing software radio platforms possessing the ability for wideband and frequency flexible operation has not yet been achieved at a reasonable cost due to existing and potential future standards.

Software defined radio is generally defined to be any radio where the physical layer characteristics are controlled by and can be reconfigured using software. To achieve this vision requires not only flexibility in the radio-frequency (RF) frontend circuits, but in the digital processing hardware as well. Progress in achieving this flexibility is being made,

The SDR Forum's definition of an ideal software radio is a radio where there is software control of a variety of modulation techniques, wide-band or narrow-band operation, communications security functions, and waveform requirements of current and evolving standards over a broad frequency range. The frequency bands covered may still be constrained at the front-end requiring a switch in the antenna system.

This paper is the result of several round-table discussion held during the First Collaborative International Software Defined Radio (CISDR 2008) Workshop, May 12-14, 2008, Maynooth, Ireland. This event was sponsored by the U.S. National Science Foundation and Science Foundation Ireland.

though slower in the RF hardware space than in the signal processing domain.

Many of the issues or constraints that are complicating the design of true software defined radios are inherent in the architecture of our existing radio designs - both hardware and software - and in the manner in which our wireless communication standards have developed over the past century. Examples of such constraints include the emphasis on robustness to strong interferers, tight filter roll-offs, rigid adherence to the OSI seven layer model, or the strong allegiance to heterodyne radio architectures. Despite the tremendous advancements in radio technology over the past century, the basic radio architectures and spectrum management rules would have been familiar to Armstrong and his colleagues from the 1930's. While there are strong arguments in supporting a conservative approach to radio design, it is becoming clear that after a century of refinement of the basic radio architecture, it is now limiting what may be possible to achieve. This becomes particularly pertinent in the coming years when resource constraints become dominant, whether with respect to available spectrum, power, multiple access, or cost.

We propose that with the advancement of our technical capabilities, both in hardware design and in signal processing, it is appropriate to reconsider the assumptions upon which we have based our existing radio architectures. It is our contention that aspects of the existing design philosophies for radio systems and related wireless standards are excessively restrictive and place undue, even unnecessary, difficulties in the path of software defined radios. A clean-slate approach may allow us to change or loosen the constraints under which we design wireless systems and thus provide us with the freedom to achieve truly reconfigurable radios at a reasonable cost.

2. ANALYSIS OF EXISTING ARCHITECTURES

Radio systems and their regulatory environment is such a large topic that no single paper could address the many relevant issues. In this section a sample of the representative issues will be illustrated and their impacts on software defined radio systems.

An important consideration when reviewing any radio architecture is to understand the context within which it is required to operate. Radio or wireless communications are normally regulated by national regulators, such as the FCC in the United States, who set constraints to which any radio wishing to utilize the electromagnetic (EM) spectrum must comply. The underlying philosophy of the regulators is relatively straightforward:

- The regulators are more interested in the EM spectrum rather than the information transmitted
- Multiple users should be able to coexist without interfering with each other

- New services that become available should not interfere with existing legacy users.

It should be noted that users of spectrum include two-way communication systems, radio-location beacons, satellite-TV-radio broadcasters, and radio astronomers. Many of these users may be defined as legacy systems with frequency allocations stretching back to the earliest days of radio. To ensure coexistence, the regulators normal approach is to require a user to ensure that any transmissions that may cross over into another user's frequency allocation be below a certain power level. Compliance is normally ensured through the use of high-order front-end filters and frequency guardbands, which creates limitations on the performance of true software-defined radios. Frontend filters are normally passive devices which, until the recent DARPA Analog Spectrum Processing program, have generally operated with a single, specific frequency range. Guardbands are deliberately unused spectrum which in an increasingly spectrum-constrained environment is undesirable, as they take up valuable spectrum real estate. Minimizing the guardband size increases the filtering requirement, leading to increased cost and reduced flexibility.

The regulators' position is that of minimal interference, however many wireless communication standards are designed with the concept of tolerance to interference in nearby frequencies while requiring excellent sensitivity to transmitted signals. A particularly challenging example is GSM where sensitivity of at least -102 dBm is required in the presence of a 0 dBm blocker. With a wideband pre-selector filter, these requirements imply a dynamic range for the radio in excess of 102 dB, which is challenging for all aspects of the receiver chain. More modern standards are less stringent but operation in the presence of strong blocking signals is a significant challenge for any software defined radio.

While the regulators and wireless standards set external constraints, radio designers utilize a limited set of architectures to build radios. The most common architectures are based on Armstrong's superheterodyne radio of 1918, typified by the presence of one or more mixers in the radio signal chain. The superheterodyne architecture has become the dominant radio design methodology due to its tenability, selectivity and amenability to low-cost implementations. It is hard to argue against such a successful approach; however the very dominance of this architecture has limited the development of other potential approaches. Over the past 100 years there have been a number of techniques that have been used and gone out of favor; for example, regenerative radios were popular in the 1910's, impulse radios were used in the 1900's and are the precursor to today's UWB architectures. It is common that technologies can disappear and later reappear as technology developments favor one approach over another. Today there

appears to be limited interest in pursuing atypical radio architectures due to lack of interest or the increasing demand to deliver radio systems in ever shorter development cycles.

Though radio frontends present some of the more intractable challenges for software defined radio, the current partition of signal processing and software elements in the radio also leads to constraints on the overall system. In the SDR/CR domains it is particularly important that the higher layers have knowledge of the physical layer performance and behavior. However, the existing models assume a strict partitioning of knowledge. These issues have led to an increasing interest in cross-layer design methodologies where such rigid partitioning is ignored and more optimal interactions are considered. An example of one issue that is particularly troublesome is the latency issue in software defined radios where many communication standards require a very fast response from the MAC (on the order of microseconds). These response rates came from a historical perspective of tightly integrated optimized components where rapid response was available and could be utilized. Many general purpose software radio systems are incapable of delivering the responsiveness required, but perhaps a more relevant question is: are these tight communication timings required or are alternatives available that might be as spectrally efficient with a more relaxed timing scheme.

The impact of these, and other, constraints is that radio design is now an optimization between the traditional size, weight and power criteria (SWAP) but also cost and flexibility. Given the constraints demanded by the communication standards and current design methodologies, adding flexibility tends to be sub-optimal choice, requiring significantly more power and cost. This is particularly true where that flexibility extends to the RF frontend circuits. Within this design paradigm, the business case for software radio remains difficult and normally depends on value-chain issues such as time-to-market or managing uncertainty.

3. CLEAN STATE RADIO

In engineering it is easy to focus on solving the immediate problem, pushing the known solution that bit forward, incrementally adding novelty. With this focus on detail it can be difficult to abstract oneself and question whether the overall approach is the optimal approach to use. In the context of radio design, we propose that it is increasingly important that the fundamental assumptions of radio engineering are challenged and assessed whether they remain valid or whether alternative solutions exist. The term "Clean Slate Radio (CSR)" reflects that concept that radio system design should be investigated with no prior assumptions. This will be challenging as many of us accept some assumptions as de-facto truths. For example: do we need a metal antenna? Do we have to support and protect legacy systems? Do we need band-select RF filters? At first

glance many of the answers are obvious, but upon reflection many of the immediate answers are based on precedence or traditional behaviour. For example, a carbon nanotube radio was recently demonstrated that used no metal or active components with opens many new possibilities [2]. The following section highlights some avenues in which traditional philosophies on radio design may be challenged.

Alternative Design Criteria

Radios, including all elements from the antenna to the network, are difficult to design as they require the combination of multiple technical specialisms in an environment of conflicting design objectives and constraints. Commonly the overriding requirement is the need to satisfy the performance requirements of a specific standard with secondary objectives such as cost, range or energy efficiency. This constrains the freedom of communications engineers to explore and develop alternative radio architectures that might provide superior behaviour in one aspect, say efficiency, at the expense of data throughput.

There are a number of trends that may provide the opportunity for increased flexibility in radio design: the proliferation of new radio applications; increasing latitude from spectrum regulators; and the ability to exceed minimum performance requirements for many existing standards. In this context, an alternative criterion may be considered as the primary objective. The following section will provide examples of some alternative criteria.

Maximize bits transported per joule of energy

There is increasing pressure, which may turn into regulatory pressure, to be more environmentally conscious. It is estimated that 3G networks tend to produce between 25 and 50 kg of CO₂ per subscriber per year (consuming approximately 160 MJ annually) [3]. A push towards the recently termed "green radio" (Hamid Aghvami) requires an alternative approach in the physical layer implementation, the operation of the MAC and the network architecture. These modifications may come at the expense of performance or reliability, for example energy use could be reduced if the network did not request re-transmits or used lower transmit powers and then corrected for increased errors through enhanced coding elsewhere in the radio.

Maximize bits transported per euro or per dollar

Cost in a radio system normally refers to the once-off cost of the equipment. For operators the recurrent costs of a radio system can be the dominant consideration, for example costs of energy, spectrum access and the opportunity cost of unserved customers. With traditional wireless devices there would be no flexibility to alter these recurrent costs, however SDR/CR technologies allow radios to be dynamically configured to optimize their costs given a users required needs or their current payment plan.

Guarantee ubiquitous connectivity.

Universal broadband service coverage is becoming an increasingly common requirement for many national operators and the opportunities presented by ubiquitous connectivity are tantalizing. Achieving this capability with existing wireless devices is proving difficult, particularly where rural, long-range, low-user density scenarios conflict with urban, dense user environments. While alternative technologies may be used in each scenario, this conflicts with the need for a mobile user to seamlessly move between regions.

Guarantee self-configuring or self-healing networks.

Dense networks of femtocell basestations, whether WiFi, WiMAX or UMTS, are proposed as a means of delivering broadband data rates to users. It is expected that there would be little or no a-priori network planning and such networks would need to detect their environment and self-configure. As an excellent example of cognitive radio technologies, this approach places severe constraints on the physical layer implementation as the radio environment will be highly congested with potentially strong interferers.

This brief review is only indicative of the alternative criteria that may guide future radio systems. In recent years there has been increased interest in previously discarded radio technologies that could offer improved performance under some conditions. Two alternative architectures that have received renewed interest are derivatives of Marconi's spark-gap generation and deForest's regenerative receiver. Spark-gap signaling could be considered an early form of ultrawideband (UWB) communications and UWB has demonstrated superior performance over traditional radio designs in certain situations. Super-regenerative radio receivers are also being explored, particularly in low power applications [4]. These, and other, radio solutions were traditionally avoided as they are not optimal for most wireless applications, but in a changing context, there is value in re-evaluating past achievements and solutions. Alternative radio circuits may offer new opportunities as can modulation and coding. It is possible to imagine complete spectrum bands managed through a single scalable OFDM plan, where users utilize one or more channels as needed. Such an approach could maximize efficiency and offer more control over spectrum usage, but would require a fundamental re-evaluation of spectrum usage.

To conclude, radio system designers are facing a future where traditional radio design methodologies will be challenged with an increasingly disparate set of objectives. Flexibility will be required and solutions may be found if commonly accepted *optimal* or *traditional* choices are challenged.

Wireless Co-Existence & The Role of Regulation

The electromagnetic (EM) spectrum is at the core of all wireless communication systems, where a transmitter can be viewed as an EM wave generator that is capable of adjusting the physical properties of its emanating waves over time in order to convey information to a receiver, i.e., EM wave receptor. Ever since Guglielmo Marconi successfully demonstrated transatlantic wireless telegraphy in 1901, wireless transmission has increasingly become an integral part of human civilization, enabling a wide range of applications ranging from financial transactions to entertainment and social networking. However, with the proliferation of wireless transmissions over the past century, especially with the advent of personal radio devices, the EM spectrum regulatory framework that has adequately served modern civilization for most of the twentieth century has started running into a *capacity brickwall* with respect to the number of supported users and amount of available transmission bandwidth. Consequently, in order for this natural resource to continue satisfying the demands of modern society, a critical rethinking of enabling more efficient allocation and use of EM spectrum is required.

In 2002, the U.S. Federal Communications Commission (FCC) issued a notice of proposed rule-making (NPRM) indicating the need to move away from the *command-and-control* EM spectrum regulatory framework currently being employed by most national governments [5]. In its place, they suggested a *dynamic spectrum access* (DSA) approach, where unlicensed (secondary) users temporarily borrow unoccupied EM spectrum from incumbent (primary) license holders. Maintaining the legacy of traditional EM spectrum allocation frameworks, these secondary users must ensure that the rights of the primary license holders are respected by eliminating any potential interference with the latter. Although this shift in the EM spectrum regulatory paradigm is considered radical by many experts from industry, academia, and the government, we believe this shift does not go far enough in order to fully exploit the technology currently available to wireless transceivers.

The primary role of any spectrum regulatory agency is to ensure that all wireless transmissions minimize, and even eliminate, any potential EM interference to other signals simultaneously operating in the vicinity, i.e., the "prevent interference to other users" model. However, this role is based on the legacy of the first radio systems, which did not possess any digital signal processing or advanced RF technology for enhancing system robustness to EM interference. However, with the advent of digital processors, microelectromechanical systems (MEMS), and other communication technologies, radio systems are capable of being highly robustness to various types of EM interference via advanced digital signal processing and digital communication algorithms. Consequently, current wireless transceivers are substantially more robustness than

their predecessors from half a century ago. Nevertheless, today's radio systems are still constrained by these same interference regulations despite their ability to adequately counteract most interference sources. To make efficient use of radio hardware resources, the interference mitigation paradigm that has governed wireless systems for a century needs to change from a "prevent interference to other users" model to a "robustness to interference from other users" model. By taking advantage of the technology available to current wireless systems, it is now possible for all radio devices to focus on mitigating the effects of the omnipresent sources of EM interference to their transmissions while simultaneously not wasting resources and time on suppressing their own EM emissions.

When radio systems were first deployed on a large scale, most transceivers focused on supporting long range transmission distances on the order of kilometers, e.g., FM radio, Radar. However, this also implies the same EM spectral band will be unavailable to other radio systems over a transmission radius of approximately the same distance. Consequently, the number of simultaneous transmissions that can be supported within a specified frequency range is constrained by the number of spectral bands available, as well as the proximity of the transmitters. Nevertheless, with the number of wireless applications and users proliferating, especially over the past decade, transmission ranges have decreased in order to enable greater *frequency reuse* of the EM spectrum. Moreover, many applications are designed to share the same spectrum via *time division duplexing*. Finally, the idea of performing *dynamic spectrum access* will eventually allow for secondary users to borrow unoccupied spectrum from primary license holders. Thus, all of these approaches for enabling greater user and bandwidth capacity in wireless access scenarios result from a better understanding of the *electrospace* by both radio planners and spectrum policy makers.

The electrospace describes the frequency, temporal, and spatial behavior of EM spectrum that is influenced by both basic EM physics and the radio propagation environment, e.g., urban valleys, rural prairie regions. Thus, future radio systems should be capable of *opportunistically* transmitting information based on the conditions of the electrospace, as well as their target signaling range. To assist in achieving this objective, the EM spectrum allocation framework should be one of an *open spectrum access pool*, where all transmissions have equal rights to the spectrum, i.e., no incumbent license holders or priority users, and that the regulatory agencies possess only a minimal role with respect to refereeing spectrum access, i.e., prevent jamming, spectrum hoarding. Moreover, these radio devices will tailor themselves to specific electrospace niches, where the transceiver configurations for urban valley and rural environments will vary due to the different associated challenges. Thus, combined with the "robustness to

interference from other users" model for EM interference mitigation, future radio systems will evolve into highly agile platforms for communications.

Impact of Cognitive Radio Technologies in Resource-Constrained Environments

With continual advances in microprocessor technology, many modern wireless transceiver systems are implementing an increasing percentage of their digital communication and digital signal processing algorithms in software rather than on application-specific integrated circuits (ASICs). Moreover, several platforms are constantly pushing the boundary between the digital and analog domains of a radio transceiver chain to as close to the antenna as possible. Finally, several platforms possess at their core a general-purpose microprocessor in order to implement many, if not all, baseband radio functions. Given this high degree of functional agility, as well as the availability of on-board microprocessing, it is possible to implement a radio platform that is capable of adapting its operating parameters in real-time based on sensory information for the transmission environment in order to achieve one or more performance goals. These types of radio systems, where the decision-making process is performed automatically and without intervention or input from the end-user, is referred to as a *cognitive radio* [6].

The ultimate goal of any cognitive radio is to take the human end-user out of the decision-making process with respect to the selection of appropriate radio operating parameters and digital processing blocks. Although several approaches have been proposed in the literature, all employ some form of machine learning implementation operating in real-time on-board the microprocessor of the cognitive radio. Current research and development efforts are focused on devising fast machine learning implementations given the time-varying nature of the electrospace, which may potentially be rapidly changing. Moreover, defining quantitative relationships between the environmental conditions and the desired operating parameters is also being pursued by experts worldwide. Thus, it is essential to have a platform that is capable of making nearly-instantaneous decisions yielding radio configurations that enhance the overall system performance and enabling adaptive algorithms that are otherwise not possible with conventional wireless systems.

Transceiver optimization is the process by which a cognitive radio attempts to (re)define its operating parameters in order to achieve some specified performance goal or collection of goals. Several common choices for goals include maximizing the overall data throughput, minimizing the transmit power, and maximizing the error robustness. Transceiver optimization is usually employed when the amount of radio resources available is limited, such as digital processing, radio battery life, and unoccupied

EM spectrum. However, in addition to selecting appropriate operating parameters, transceiver optimization techniques can also be used to decide on an adequate set of functional blocks that can be employed by the cognitive radio platform. Consequently, choices can be made regarding which block to include in an implementation and which to omit due to resource constraints and redundancy. For example, a transceiver optimization routine would need to decide whether error correction codes should be employed when adaptive modulation is already available to the system. In other words, the transceiver optimization routine would assess the *added value* of each block to the overall performance gain of the system. As a result, quantitative trade-off analyses between different options that may or may not appear to be independent are required. Given the potential limitations of most radio platforms with respect to available processing power and memory, power supply, and implementation costs, getting the most out of the platform for a specific set of goals while simultaneously balancing these trade-offs is possible with cognitive radio.

Impact of Network Architectures on Radio Design

Traditionally wireless systems were defined broadly in line with the OSI 7-layer model with strongly delineated boundaries between the physical, data-link and network layers. This layered, or modular, approach facilitates the rapid development of new services as developers in each layer able to develop optimal solutions. In practice this separation of roles was rarely as clear-cut as suggested in theory and the development of SDR and cognitive radio technologies has blurred the boundaries further. From an architectural perspective the separation of functions is increasingly false – the choice of network structure has a significant impact on the radio environment in which the device must operate, and the choice of application (eg voice versus streaming video) has impacts at all layers. For cognitive radios all layers must co-operate to select and implement appropriate policies.

It is also important to note that the 7-layer approach is sub-optimal, both in implementation complexity and link performance. There is information available in the physical layer that could assist in quality of service or energy consumption that cannot be used. This has led to a growing interest in cross-layer design methodologies where physical layer data is passed to the higher layers and aspects of the various layers (for example routing or source coding) are modified to optimized various criteria [7]. This can yield benefits for quality of service, data throughput and reduced energy consumptions. Unfortunately cross-layer designs are not always beneficial and can lead to complex interactions between the layers due to the various feedback paths.

In the context of “clean-slate-radio”, the OSI model of separated layers should be initially discarded and only adopted where beneficial. The FIND initiative, working on

clean-slate-Internet” has proposed alternative architectures for designing radio networks. Others have proposed radically different radio receiver architectures, for example direct analog to channel symbol conversion through Shannon mappings. In summary, the separation of functions in layers has significant benefits but is not a fundamental feature and, where a case can be made, should be considered optional and ignored if necessary.

4. SUMMARY

The objective of this paper is to suggest that there are avenues and opportunities to reevaluate the underlying assumptions in our radio design philosophies. The concept of a “*clean slate radio*” is that one should identify those aspects or features that present difficulties or challenges, and examine whether the causes of these difficulties remain valid in all scenarios. With the rapid development of science and engineering, new radio technologies are being developed and older, discard schemes, can become relevant again. It may be possible that all our assumptions are optimal, but it is always beneficial for researchers to challenge our assumptions and to be open to a revolutionary approach to radio design especially in light of increasingly disparate and non-traditional requirements.

5. ACKNOWLEDGEMENTS

This paper is as a result of the NSF International Workshop on Next Generation Open Architectures for Software-defined Radio held in the National University of Ireland Maynooth in May 2008. The support of the National Science Foundation (NSF) and Science Foundation Ireland (SFI) is greatly appreciated.

6. REFERENCES

- [1] J Mitola, "The Software Radio," *IEEE National Telesystems Conference*, 1992
- [2] K. Jensen, I. Weldon, H. Garcia, and A. Zettl, *Nano Letters* pp. 3508-3511, 2007; corrections: pg 374, 2008.
- [3] "Sustainable energy use in mobile communications", *white paper, Reference EAB-07:021801 UEN*, Ericsson, 2007
- [4] FX Moncunill-Geniz, P Palà-Schönwälder, O Mas-Casals, "A Generic Approach to the Theory of Superregenerative Reception", *IEEE Transactions on Circuits and Systems I*, pp 54-70, 2005.
- [5] Federal Communications Commission, "Spectrum policy task force report", ET Docket No. 02-135, 2002.
- [6] J. Mitola, "Cognitive Radio for Flexible Mobile Multimedia Communications," in *Proceedings of the IEEE International Workshop of Mobile Multimedia Communications*, vol. 1, (San Diego, CA, USA), pp. 3–10, November 1999.
- [7] Srivastava, V. Motani, M., "Cross-layer design: a survey and the road ahead", *IEEE Communications Magazine*, pp. 112-119, 2005.