Deep Science Publishing
https://doi.org/10.70593/978-81-989050-5-5_4

Chapter 4: Maximizing wireless bandwidth efficiency through spectrum allocation and cognitive transmission

4.1. Introduction

In traditional wireless communication systems, a limited amount of spectrum is allocated for each service. With the accelerating growth of wireless data traffic, major Internet service providers in the United States have been upgrading their backhaul connections to 1 Gbps fiber-optic lines, while the wireless last-hop link has become a bottleneck achieving only 500 kbps due to the limited amount of allocated spectrum. Because the actual physical bandwidths of wireless links at frequencies lower than 10 GHz are orders of magnitude larger than that of 500 kbps, only a small fraction of the wireless physical bandwidths are utilized for communication. Maximizing the bandwidth efficiency of wireless communication systems means that maximum user capacity can be served for any given amount of spectrum allocation since the user capacity is also directly proportional to the number of code-books available to each user (Brown & Iqbal, 2025; Chen & Roberts, 2025; Kumar & Duong, 2025).

This requirement of using only a small amount of allocated spectrum is fulfilled by the use of a simple non-directional antenna together with either time-hopping or direct-sequence spread-spectrum signaling. Such communication systems experience a very large number of simultaneous interfering users and rely on the modulation-code-collection strategy for establishing connections between a transmitter and a receiver among the large number of users. Such methods of connecting long-distance transmitters and receivers have been practiced by the cellular telephone industry without micro-wave antennas or spread-spectrum signaling. On the other hand, securing a large physical bandwidth in the presence of a large number of simultaneous users with small allocated bandwidth means working with a large, and yet of equivalent dimension size, wireless branch of the electromagnetic spectrum (Kumar & Duong, 2025; Perez & Hassan, 2025; Reddy & Lin, 2025).

4.1.1. Background and Significance

In the last decades vertical and horizontal contextualizations let wireless communication evolve remarkably. In the vertical aspect, wireless communication has to serve large and small users in different access and transport for voice, Internet, data, gaming, video, multimedia services and particular affluences and guarantees in Jitter, delay, and lost packets. In the horizontal aspect, wireless communication had to support not only cellular but also hot spotting and local area communications in the university, corporate, café, and hotel environments. These efforts are drafted into standards.

Going forward, the convergence of telecommunications and information technology supported by the high-speed wireless backbone and based on the scales of incorporation of digital and wireless technologies is splitting the telecommunications segment and globalizing the scope of the technology. Indeed, the separation between the telecommunications and computer industries is increasingly becoming blurred. With the convergence of the two sectors, the expertise, customer base, and distribution network of telecommunications and computer companies may move towards each other. We may witness computer companies moving toward coaxial and fiber-optic modulation for domestic telecommunications, and traditional telecommunications firms:

• fusing their telecommunications networking technology and infrastructure into high network performance servers for scaling the digital economy;

• moving further through the integration of terrestrial and satellite wireless backbone technologies toward mobile cellular transport.

Wireless communications may continue to grow larger not only in sustaining the dynamics of traditional self-contained business segments uplifted by user-friendliness demand but also in globalizing the conduct of the business. In addition to standards, international disciplines, cultural acceptance, technology, and cross-border investment need to evolve further. The discipline and motivation for this evolution happen to be, no surprise here, large economic and technological scales.

4.2. Understanding Wireless Bandwidth

To better gauge how wireless bandwidth is produced and controlled, it is helpful to step back and go over some of the basics of what wireless bandwidth is, who controls it and how it is controlled, what factors matter in these controls and assessments, and what their limitations are. To be more specific, with time, people have become increasingly aware of wireless bandwidth. It is taken for granted, and the focus is ultimately on the products and services that it allows. Take for example radio broadcasting. It has been the same programs running uninterruptedly on air, attracting six or seven decades of sharing laughter and tears among parents and their kids, and keeping the family together in an era long before the advent of television, with all its various formats. Similarly for point-to-point wireless connections. Take for example local air traffic control, radio communications, and cell phone communications. Be it for safety, security, corporate telephone systems, or personal conversations, the same wireless airwaves have been occupied repeatedly and forever. But at what point is the amount of wireless bandwidth consumed worth celebrating?

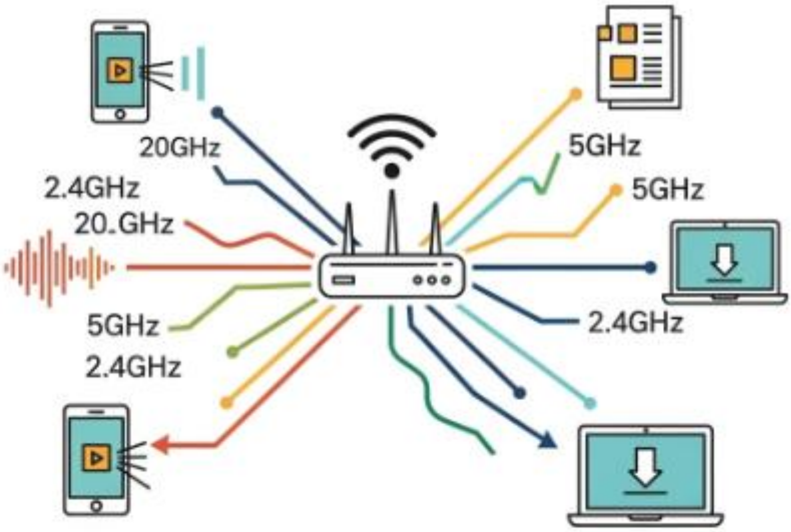

Fig 4.1: Understanding Wireless Bandwidth.

Wireless bandwidth has become a needed element of business success that was once the envy of TV and movie producers, local and international air traffic controllers, corporate telephone systems, the military, and the federal government. But wireless bandwidth never actually used to belong to anyone, be they corporations or countries. Actually, it is like air, which everyone consumes and pollutes in various forms as they go about their daily living. What is different for wireless bandwidth is that it must be carefully managed, consumed, and protected so that everyone can share in it and nobody can monopolize or disrupt it. Furthermore, these levels of management and protection for wireless bandwidth are assigned by regions, nations, and international bodies. These controls affect how wireline and wireless bandwidths are determined within the broader context of communication networks. It is these resource levels of wireless bandwidth that must continually be maintained in conceived and designed networks to meet the diverse communication and commerce needs of industries, companies, and social groups.

4.2.1. Research Design

This research was conducted as a qualitative, phenomenological case study to investigate the phenomenon of bandwidth shortage and how obstacles preventing efficient solutions for the increasing demand for communication services were perceived by a cohort of ten industry practitioners. Working in wireless spectrum management organizations, these practitioners together brought over 180 years of experience in the telecommunications industry, almost equally disseminated between top management, middle management, and operational functions within ten government agencies, commercial companies, and research organizations across eight countries. Over a 7-month period, semi-structured interviews were conducted with the cohort to both seek conceptual clarification of the participants' perceptions and to triangulate findings through within-case and cross-case collection and analysis of other data relating to the cohort's perceptions. Computer-aided thematic analysis software was used to support the construction of linear narratives and code development.

These narratives were developed to reflect the cohort's perceptions both during and following the main phase of interviews. The longitudinal design part way through writing the narratives was adopted to explore embedded time changes in the cohort's longitudinal changes in perceptions. The cohort was then queried to test the contextual validity of the narratives. The cyclical approach to data collection, analysis, and member checking was continued until thematic saturation between the longitudinal narratives and the investigators' previous knowledge was reached. The output of the study, coherent and accurate accounts describing the meaning of the phenomenon for a small group of practitioners over time, were informed by the latter's reflections during successive longitudinal interviews, personal documents, and semi-structured email correspondence with the cohort. Their findings were subsequently contrasted with the findings from traditional bandwidth research, comprising the quantitative submissions to peer-reviewed journals, books, and doctoral dissertations.

4.3. Spectrum Allocation Strategies

Unlike the conventional and classical networks, in wireless networks, the bandwidth is a scarce and non-renewable resource which, if not allocated efficiently among users, can lead to conflicts and ambiguities. Many researchers believe that the key question in successful use of mobile communication resources and services is allocation of bandwidths in space and in time, that is, spectrum allocation. Bandwidth efficiency is defined as the ratio of a measure of performance to the total bandwidth allocated. A variety of approaches to the problem of allocation of bandwidth have been proposed over the years. Bandwidths may be set for a given link or route which would then remain available for use by that link or route over a considerable time. Such static assignment of bandwidth would be feasible for such links or routes for which persistent custom exists. A more flexible alternative is dynamic assignment of bandwidth that is adapted to the instantaneous condition of the various links. The bandwidth allocated to a link would vary from one allocation period to the next, depending on the state of the link. Such a technique of dynamic allocation leads to an efficient use of system resources. At the same time, it does have operational difficulties especially with regard to signalling.

The practical problem is of assigning diffusing resource such as bandwidth at a specific time over a specific period in what is today assumed to be a radio fixed network. Given this condition, when does the network manager switch from one channel to another? What is the choice of channel? What is the effect of multiple access on the switching activity? How long should a call be allowed to remain in a blocked state before changing the choice? These are just a few of the many questions that need to be answered before a capacity sharing scheme can be successfully implemented. In this section, we present and discuss various strategies for spectrum allocation which can be static or dynamic.

4.3.1. Static Spectrum Allocation

Spectrum allocation is mainly divided into static and dynamic schemes, according to the assignment time of the parameters introduced by designers. The basic concept of static allocation is as follows: the system decides the bandwidth and the allocation of each band in the initial design according to the predicted traffic load in a certain period of time in the future. Many systems and services for public use are designed for long-term stable use. These systems provide large single-channel capacity, such as by assigning wideband to local radio and television programs. However, spectrum demand varies greatly with time. For example, the demand for radio mobile frequency during the day is much greater than that at night. Because static allocation cannot meet the time-varying characteristics of user requirements, there are often some frequency bands that use relatively little time. In addition, the short-wave system provides communication services for shipping, aviation, and other transportation sectors, basically only at special hours.

Static allocation also has many other technical restrictions. For example, some radio authentication technical parameters of various kinds of radio and television stations need to be matched, or special frequency equipment for countries that are far apart cannot be approved. A static allocation plan is given in the frequency table, listing multiple bands used daily and nightly for these communications. The course of these radio services adopts many different patterns. The allocated irregular band for individual land mobile radio service provides service during a period of the day or year. In addition, some types of regional service radio systems provide communication services according to the regular opening and closing schedule of communication circuits, or take the allocated irregular frequency or time slots according to the demand, and adjust the starting or ending time accordingly. Electronically controlled scanning systems could also apply the demanded time of assigned channels of regional service according to the time on which the system switches assignment patterns.

4.3.2. Dynamic Spectrum Allocation

In dynamic spectrum allocation, multiple spectrum piles exist simultaneously, and the spectrum in the pile can be used at different times. To support fairness and increase the system throughput, both primary and secondary users are allowed to transmit simultaneously, but the sum data rates should not violate the tolerable interference level by each user. For periods with no primary user transmission, the frequencies can be allocated to either primary or secondary users. We observe that the diversity in time or frequency can increase spectral efficiency. Dynamic spectrum allocation is generally acknowledged to be a highly effective means to alleviate spectrum scarcity. Many results in game theory and economics have shown that free markets can efficiently allocate spectrum for different types of users, and that different types of users can cooperatively maximize the value of spectrum licenses. In DSA, secondary users are allowed to access parts of the spectrum primarily held by other users, but they have to protect the primary users from harmful interference.

Dynamic spectrum access is capable of improving spectral efficiency significantly. Moreover, systems can operate in a mixed state - both temporally fixed slot assignment for the primary system and dynamic access for more non-disruptive secondary systems. Both primary and secondary users are allowed to negotiate transient spectrum ownership. Dynamic spectrum access removes temporal access inefficiencies, and increases the value of the spectrum resources by permitting more users to share the resources of selective pathways during short time durations. DSA has a number of socioeconomic advantages over primary usage systems. Many primary and secondary users can gain real-time access to hotspots of very high, but time-limited values. Systems can utilize high-capacity hotspots for short periods of time at much lower fees than would charge for full-time access.

4.3.3. Spectrum Sharing Techniques

To enhance spectrum efficiency, a number of techniques that allow secondary, unlicensed users to share a portion of the licensed user spectrum, while causing no interference to the primary users, may be exploited. The basic idea is to allow unlicensed users to transmit for limited periods of time, requesting as little service time in response to their requests for service as possible, to meet their QoS requirements. Heterogeneous environments may also be explored, where the primary system service requirements are accommodated in order to create greater availability for unlicensed users, for example by employing carrier power balancing methods. Spectrum on demand service models may also be utilized, where the primary system temporarily gives up control of the spectrum, requesting an instant channel hold up time for the secondary service requests, while the primary user has been warned to occupy only a short time span in the immediate future, again with potential allowable QoS violation penalties associated with sporadic primary system hold on time violations.

The first class of techniques that appears is the so-called inactive slot or ready stream approach, where the primary system random bit stream is retransmitted by the secondary unlicensed system on the unused bit slots, causing no interference to the primary licensed system while it is in transmission. The advantage of this approach is that at peak load times, the capacity of the primary licensed system could be inhibited, since the capacity of the unlicensed system is zero. The major difficulty is that any kind of fairness service model among both systems with all the drawbacks associated with it is rather hard to implement, given the nature of the primary system behavior, and that the only way in which some fairness might be obtained at the secondary system level would be through complex signaling and feedback mechanisms.

4.4. Cognitive Radio Technology

We are immersed in a wireless ecosystem in which every day more users demand diverse services which, when not fulfilled, negatively affect our lifestyle and activities. However, there is a paradoxical situation: there is an abundance of available radio frequency spectra, but the regulation and allocation of this resource is so complicated, and often so inefficient, that it ends up being underutilized – despite the huge demands. The idea is simple: allow unlicensed users to use the parts of the spectrum which are currently not in use by the licensed holders.

Since costs are much less than what it would take to build the whole infrastructure, it is becoming increasingly feasible to share what is already there. One of the technologies enabling such shared usage of the wireless medium is Cognitive Radio. In general, Cognitive Radio is a term which applies to opportunistic, unplanned or ad hoc wireless networks of uncertain topology. In particular, Cognitive Radio describes a network in which participating devices have the capability of sensing their environment and establishing methods of communication given limited prior knowledge – parameters such as frequency hopping sequence and modulation type are discovered or refined during the short duration of initialization of link setup. While providing internal intelligence in the design allows smart collaboration, without central coordination the movement of users can create safety concerns among different groups that may be sharing space, such flexible systems are inevitable as we reach for the efficiency possible with large networks.

4.4.1. Definition and Principles

Cognitive radio (CR) technology is one of the key components for improving wireless bandwidth efficiency. The cognitive radio (CR) is an intelligent, computerized wireless node that learns from its surroundings. It senses the frequency spectrum around it, learns from its previous interactions and those of other radios, and adapts its RF parameters for optimal operation. The combination of its sensing capability and the knowledge it acquires, CR can change its operating wavelengths, transmission timing and power level to minimize interference into other users or user groups. In this definition, CR is considered a software-defined radio (SDR) with specialized intelligent detection and decision-making capability. The key features of CR technology, as compared with conventional SDR technology, are incorporated in the following four principles. The first principle is "causing no harmful interference to other users". An important feature of the CR is that it does not actively participate in the communication, but functions only in the sensing and learning mode until it learns the spectrum usage pattern of other radios in its vicinity. The second principle is "having the user programmable capabilities". In an SDR, a user can download any radio operational software remotely. This enables the SDR to become a CR for any other radio, thus turning clusters of SDRs into a mesh community of cognitive radios. The third principle is "learning from its previous interactions". The CR learns from having sensed the RF spectrum and the timing of activity of other radios in its vicinity. The last principle is "making cooperative decisions". The CR exchanges its learned spectrum usage knowledge or its current sensing results with other neighboring CRs utilizing an ad hoc network or a software-defined mesh network. Based on this cooperative knowledge, the individual CRs make their independent decisions.

4.4.2. Cognitive Radio Architecture

A randomized, user-centered approach is adopted during segmentation as a method for controlling user bias during exploration. The basic functionality of cognitive radio requires the support of intuitive design concepts that meet the requirement for non-intrusiveness regarding the primary user. A general concept is that the secondary user wants to utilize the frequency band of the primary user when it is not being utilized, but the secondary user is not aware of when the primary user is transmitting or through what protocol.

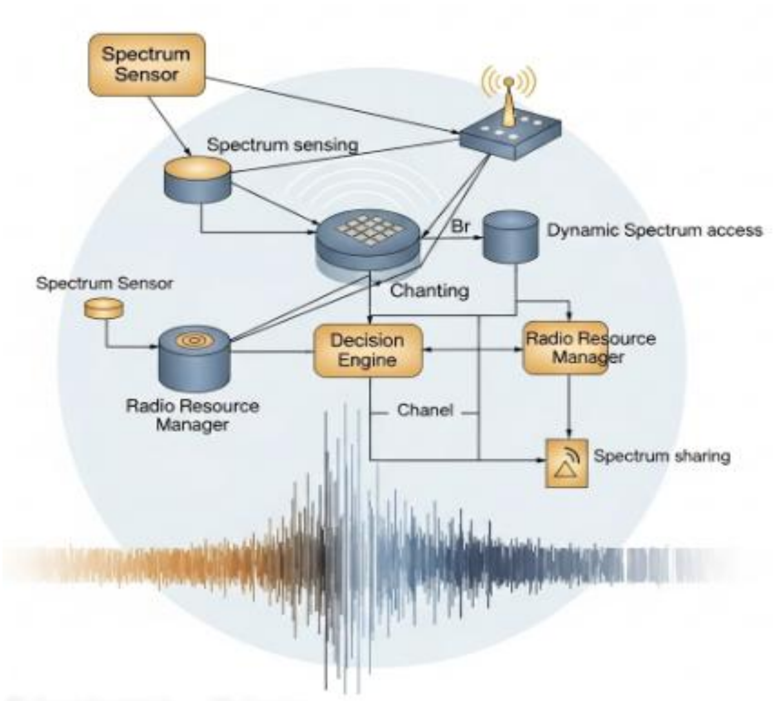

Fig 4.2 : Cognitive Radio Architecture.

Thus, a broad coverage area of the primary user is maintained for preventive and reactive interception of attacks in the primary band. Therefore, the concept of a smart, genius, or shadow user can be adopted conceptually. This perspective presents challenges regarding when the channel is being observed passively, cognitively, and/or dynamically. Moreover, cognitive cycles can be introduced during the reception timing of the user. The supporting architecture is conceptualized at the high level in which the different cognitive cycles occur. The rapid cycle handles the frames and packets and consists of the architecture used for packet-based communications. The medium cycle handles the bursts and is used for burst communications. The slow cycle handles the flows and is used for voice-based as well as real-time communications. Within each of these three different layers, multiple sublayers can be defined. The cycles within a layer may be at the same or different positions.The new concept shown previously is the resource slice in vertical and horizontal directions, each of which manages the collocated wireless devices. The slice is treated as an intelligent entity that generates or receives actions that are based on group data that is augmented with personal data. The implementation of the entity is achieved during the dynamics of the cognitive cycles. There are two basic solutions to implement surface slices: Through the superposition of signals and coding words; and through the surface waves or surface coils. Waves propagate through different partial phases featuring the surface slices superimposed by surface waves or coils. The proposed architecture is conceptualized based on these solutions.

4.4.3. Applications of Cognitive Radio

The cognitive cycle can be applied to different radio applications. This allows for different levels of transmission intelligence depending on the actual application. Current applications can be classified into four main categories. Although higher cognitive capabilities may not be required or possible in all cases, the cognitive cycle can help in different ways: reduce signaling overhead and delay, optimize the occupancy of the frequency band and minimize interference with incumbent users.

Besides SA, SC is important for DR. The problem of SA and SC is then defining how to optimally allocate the scarce resources depending on the current amount of traffic and channel conditions. The classic way to do that is to fix the scheme parameters based on statistical information, such as the average traffic load. When traffic is sporadic, both PU and SU must suffer overhead and delays until the resource is released. Currently, mapping the various traffic classes to static priority queues is the main implementation on the market. Implementing cognitive capabilities as part of the radio technology will significantly enhance the performance of cognitive networks supporting the coexistence of SUs and PUs by allowing more flexible policies in resource utilization.

In some applications, a user detects the presence of other users without establishing a set of rules for resource allocation. Nevertheless, to share the same resources without causing contention or delay or error rate degradation, the user has to have some level of cognition. The actual implementation of TRC would allow CR to make very efficient decisions and, for example, have reliable secondary access to frequencies assigned for primary use. But such TRC would go beyond the current capabilities of CR. New transmission technologies have designated cluster network structures over which wireless communication is made possible.

4.5. Challenges in Spectrum Allocation

Accessing the spectrum resource has traditionally been handled by regulating authorities around the world in a static and centralized manner. Such allocation is not without consequences, as it has led to a low degree of activity in many of the bands that have been allocated. This incoherence of demand and supply has resulted in an economy marked by shortage in some regions and considerable waste in others. Demand-induced limitations on economic growth drive economies to pursue solutions to maximize efficiency while minimizing restrictions that unduly limit growth capability. In response, during the past several years, authorities have been reviewing the current allocation of bands and are revising these allocations. Two significant results of this review are: (1) The need for more segmenting of frequency space into smaller and more precisely tailored segments that can more efficiently support specific applications. (2) Providing increasing incentives for unlicensed shared usage of specific bands.

Because of the wide diversity of applications that telecommunications systems must support, many distinct types of market leads to demand and many more leading to demand-induced challenges are present. In this chapter, the focus is on describing and explaining some of the challenges in allocation of bandwidth limitations that directly affect users and service suppliers. These challenges arise with the introduction of new users operating at different levels of quality and access cost in the same geographical location as carriers or service suppliers. This puts into question the conventional wisdom governing resource allocation. It is, therefore, an important problem to examine the methods governing demand-induced limitations. These methods relate to the management of interactions and conflicts between competing demand for bandwidth.

4.5.1. Interference Management

Spectrum allocation is the problem of allocating an incompatible range or ranges of frequencies to clients wanting to transmit in the wireless medium. Incompatible means that simultaneous or nearby transmissions from different clients on the allocated ranges will result in interference. The conventional usage of spectrum allocation primarily assumed the allocation process was in the domain of some regulatory authority. An area, which was usually the domain of some special purpose local regulatory structure, would be examining the requests of clients guaranteed to provide interference-free and adjacent transmission through the use of electromagnetic hybrids, very high bandgaps, and absorptive material, and allocating the frequency range to them. Modern applications, however, as more and more clients have to be served have exposed the inadequacy of this approach. A major drawback is the time taken to process requests, especially on short notice: Many radar systems will typically desire to use a band for a few msec every few minutes.

Another drawback is that such a centralized mechanism is clearly subject to the vagaries of specific locations and environmental conditions-precluding lower-cost, lower-power clients such as desktop computers and walkman-style receivers. For this reason, modern wireless networks have begun to prefer distributed algorithms that could provide error guarantees. However, they have other types of performance shortfalls such as assured secrecy, multiple frequency errors, multiuser detection, and overhead in terms of successfully reusing the channels in a typical bursty traffic environment. These factors have stimulated a search for more explicit and somewhat centralized distributed algorithms, such as interference cancellation. Such methods are currently being used only on a very small scale but network operators worldwide are trending towards a generalized short-range integrated wireless services.

4.5.2. Regulatory Constraints

Regulatory authority has a large number of primary users who have been allocated significant wireless bandwidth portions for a number of reasons. These reasons range from long government and military decision times to the need for planned interference mitigation to ensure link performance over geographic areas due to receivers with very good sensitivity enabling link performance in low intercepted signal to noise ratio regions. These primary user allocations based on a select set of supported service types present a considerable challenge for the allowance of secondary users in the space between these allocations. The traditional reason for this partitioning is that regulators desire primary users to get high performance and that they have a very limited amount of wireless resources.Further, the reasons for the regulations are political rather than need based. For example, there may be eight allocations for weather satellites, where only two may be used at any particular time. These unused allocations may be geographically displaced at any one time and/or frequency displaced at different times. Internationally, large portions of wireless frequencies have been partitioned for users in other countries. To mitigate interference with primary users during the time they are using their allocated frequency, primary user-specific databases will be required. It is not possible to expect primary users to put up with continuously bad service. Further, the service-level negotiations between the secondary provider and the primary user may be very complex and time consuming. In particular, spectrum requests during the winter holidays for both the northern and southern hemispheres may become important factors due to the concentration of weather satellite use.

4.5.3. Market Dynamics

In this section, we discuss the economic constraints surrounding classical policy approaches to multiple spectrum owners. These complications stem from economic literature clarifying how property rights and incentive structures drive user behavior, and production engineer literature pointing out that without an assignment process, and need for repeated stages of allocation at rapid time-scales, use of a scarcity state is not needed in a centralized approach. The complications arise mainly from the centralized approach having no corrective feedback mechanism.

It could be that the relevant authorities are not unaware of the aforementioned complication, and further, that the regulatory body is only carrying on with centralized architecture to sustain a traditionalist market view. A non-centralized architecture permitting market forces to determine good-time and bad-time prices for frequency slots, along the lines of budgeted design, then permits our collaborator to engage in maximum utilization and self-correcting price levies at bad times that induce others to resolve conflicts in different ways. In other words, under an architecture designed to discover "creative destruction", our collaborator can do the utmost to ensure that goods brokered on her behalf are being delivered efficiently.

That said, she can use the relevant situation to induce developers to add capacity on the structure, where they can boost net growth. If necessary, recovery of wireless growth and performance can also be aided through semi-centralized cooperation. So in that sense, market dynamics encourages capacity additions in both architectures. However, for market types with many participants, the decentralized architecture should provide more insights to decisions. We conclude this section remarking that engineering/service design pushes quickly to adopt a decentralized architecture.

4.6. Bandwidth Efficiency Metrics

The previous section discussed the factors affecting wireless network throughput. A key assumption of the section is that output throughput is a valid measure of bandwidth use efficiency. In other words, when wireless system operation is described in terms of throughput, and input bandwidth is assumed given by the maximum channel signal-to-noise ratio product, a simple expression of bandwidth efficiency results. This leads to the following definition of bandwidth use efficiency.

Where is the probability density function of the input channel SNR. Note that in a multi-channel system, is the sum of the PDFs over the supported channels. From this definition, if the term turns to zero, this means that the bandwidth use is zero. This term can be used to evaluate if the communication system is well adapted to use the entire channel bandwidth. In this case, because the discriminability at the output passes through supremum, this means that we could use a more compact basis at the input, which could contribute to a better waterfilling distribution.

To be precise, throughput measures input-output relation for packet type transmission systems, though other point measurement definitions exist, they are usually mere variations of the throughput definition. Systems that do not directly operate upon packet size data, such as digital speech scramblers, motion picture codecs or modulation systems, but that are able to present their performance results in terms of packet transmission throughputs, are not included in this definition. By packet type transmission, we mean a service in which the message consists of a finite number of bits, and the probability of bit error is very small and remains constant at least in the time frame required for transmission.

4.6.1. Throughput Measurement

Throughput is one of the most basic, at the same time most important, performance characteristics of a wireless link. If we ignore the implementation of Quality of Service, giving a good micro network behavior, the user's experience is driven only by the throughput measured by the user application task. Hence, no other metric could be richer in information than a bandwidth measurement, in any given statistical distribution, in bipolar circumstances, say a busy period followed by a few idle periods, paying attention to the "pay it forward rule", i.e., how a user application reuses the previously cached data. This simple rule is sufficient to warn us of abuses, where non-existing data are browsed, or the thread of newspaper pages moving fast.

Consequently, efforts need to be given to delivering a good statistical estimate of a single throughput measurement session. Here "single" means, at least, the click-oriented structure imposed by the application measurement task. Processing little samples is not reliable of a general behavior, and considering only the user task each time and holding the bottleneck congestion probability is useless either, so our understanding is standing still in the paradoxical discussion of congestion control: "it would be more convenient to reduce the process into classes according to various user-defined metrics, or observe the behavior of few connections for a given amount of time, and argue that a statistic based on input/output parameters can describe the satisfactory behavior of a network."

4.6.2. Latency Considerations

Some applications use data packets in a transaction mode, such as sending an email or transferring a file. These types of applications are sensitive to delay on a larger timescale. Other applications rely on a continuous stream of packets, such as teleconferencing and voice over IP. These applications have to abide by stricter delay and jitter constraints on a shorter timescale. Some applications, such as web browsing or online gaming, use a combination of both types of traffic where visual information is periodically requested for download from a web server or the server has to be actively kept updated with the client's state. Such applications may benefit from delays, jitter constraints, and bandwidth guarantees that are tighter than those only for transaction-style applications but looser than the minimum requirements.

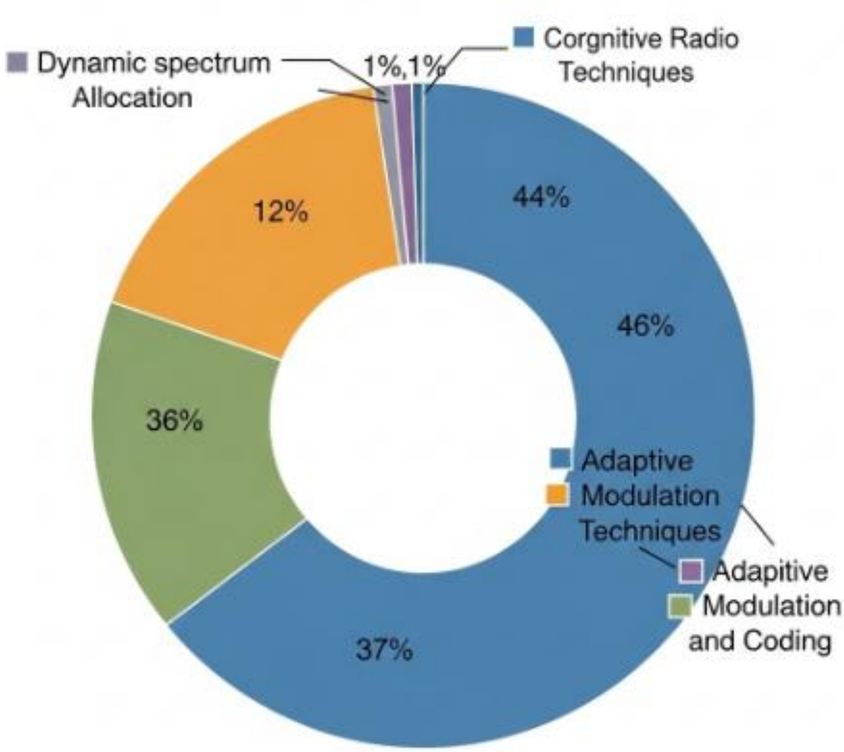

Fig : Maximizing Wireless Bandwidth Efficiency through Spectrum Allocation and Cognitive Transmission.

In terms of continuous streams of packets, the latency requirements may be broken down into three other metrics, as applicable in addition to the minimum delay. These additional metrics are (1) maximum delay; (2) delay jitter; and (3) delivery assurance. In the context of this discussion, the delay is the time taken by a packet to traverse from source to sink. The maximum delay requirement is due to the fact that voice and videoconferencing packets have to arrive within certain timeframes to constitute an understandable conversation by humans, which is usually in the order of milliseconds or a few seconds. The delay jitter requirement needs to be met since certain applications may require a continuous stream of video or voice data. If the delivery time varies drastically, the resulting audio or video may appear choppy or as if it is skipping frames. The delivery assurance constraint is important for voice and videoconferencing traffic. Since these applications are extremely delay-sensitive, the loss of low-bandwidth packets cannot be compensated by delivery at a later instant.

4.6.3. Quality of Service (QoS) Indicators

Especially for timesensitive applications, it is not sufficient to measure throughput and latency only. Their values may be excellent, but the application may still complain of bad service. This is quantified in terms of the loss rate or more generally in terms of a QoS metric. This provides the possibility of taking into account the heterogeneously variable nature of applications such as voice, video, and data. QoS is an abstract characterization that can be made specific for particular classes of flows by assigning them a value of a very particular network performance metric. Quality of service generally comprises individualized requirements regarding the various performance metrics. Thus, it is very application specific and even flow specific. The end goals of QoS research are to formally specify the various parameters that constitute a QoS metric and develop innovative and efficient algorithms for the classifications, reservations, allocations, and scheduling of resources to flows as per the QoS they require. In this, there are several areas of open research problems. No effort, however, is being spared at providing designs of the architecture and algorithms to support QoS in next-generation broadband communication networks. In the basic client-server model of bandwidth management in traditional data networks, the server provides resources to try to improve the average case performance and thereby much of the time is wasted. The specialization of the QoS metrics for particular classes of flows presents scope for modeling for more efficient allocation rules in the use of high bandwidth. QoS management particularly for delaysensitive multimedia applications requires the adoption of the idea of virtual connections with a certain bandwidth guarantee allocated to them, along with soft connection requirements.

4.7. Conclusion

These pages focused on the description of mechanisms that maximize the bandwidth usage efficiency for wireless communications. To this aim, we considered two main issues: the spectrum resource allocation and the techniques used for data transmission on the physical channels. For the spectrum resource allocation, we focused on procedures allowing the dynamic allocation of the spectrum channels with the aim of maximizing the network throughput. Such procedures are made possible by the presence in the wireless network of cognitive nodes able to actively monitor the operating conditions of the available channels and dynamically relay messages even on a channel which is not allocated to them. While channel allocation, routing, and the establishment of end-to-end paths are performed in the network layer for wired networks, in wireless networks, they are performed in the MAC because of the different scenarios of both networks. Transport layer protocols, typical for wired networks, need to be carefully redesigned for wireless environments, and the specific characteristics of wireless communications need to be taken into account. Unfortunately, this is still an open issue, and many questions remain. How to avoid too frequent connections' disconnection?

When to close a connection? How to signal for a connection closure? For the sake of power savings, how many nodes need to be in sleep mode? How to guarantee Quality of Service for sensitive applications? These questions and more remain to be properly addressed.

To this aim, it is expected that new generation cognitive wireless networks, with their pervasive nature and inclusion of nodes with very different resources and power capabilities, will create novel challenges to both academia and industry. Tomorrow's research will also focus on cognitive radio concepts for the nano-scale, and on cognitive wireless body area networks, for instance, for health-care monitoring applications, which will require the widest support of multidisciplinary skills.

4.7.1. Future Trends

Wireless communication systems have been developed during the past 100 years and have seen remarkable progress. During the last 20 years, they have spread rapidly both in terms of geographical coverage and in terms of accessibility and number of users. Many innovative techniques have been developed to address the severe bandwidth shortage problem, and network and resource management strategies have contributed greatly to the realization of this remarkable success. However, despite these accomplishments, the limitations of the wireless communication systems still persist and indeed have become a bottleneck for the emerging wireless services and applications. Perhaps the biggest implication of this work is that our wireless systems are currently focused on maximizing an abstract utility function based on traditional telecommunication technologies and methodologies. New and innovative wireless designs will not be based on marginal improvements of our current systems, but will rather shift the networking architecture towards new, more efficient solutions.

These revolutionary designs will address our wireless infrastructures and wireless networking software, with a goal to achieve intelligent allocation of the wireless resources, such as network routes, operating frequencies and bandwidths, space-time-coding, and transmission power. As our telecommunications move further away from the traditional architectures, the interfaces with other systems will take on increasing importance. Finally, we have yet to apply many intelligent techniques inspired by more advanced technological fields to our wireless infrastructures.

References

Kumar, P., & Duong, L. (2025). Architectures for AI-driven spectrum sensing. Sensors, 25(3), 1442.

Perez, D., & Hassan, R. (2025). Evolutionary algorithms in next-gen wireless chipsets. ACM Transactions on Embedded Computing Systems, 24(2), 1–20.

Chen, F., & Roberts, S. (2025). Low-noise amplifier innovations for THz communications. IEEE Microwave and Wireless Components Letters, 35(1), 56–59.

Reddy, S., & Lin, Y. (2025). Smart grid-aware communication hardware for IoT. Energy and Buildings, 310, 113573.

Brown, T., & Iqbal, M. (2025). Intelligent thermal management in multi-core smart circuits. IEEE Transactions on Components, Packaging and Manufacturing Technology, 15(2), 188–197.