Effective Access to Radio Spectrum

Tricia J. Willink
Communications Research Centre
3701 Carling Ave., Box 11490, Station H
Ottawa, ON, K2H 8S2
CANADA

email: tricia.willink@crc.gc.ca

ABSTRACT

Congestion in the spectrum bands used by military applications is limiting the use of current capabilities and the introduction of new ones. The inflexibility of the current spectrum management method, which relies on careful and complex planning to provide spectrum access for a wide variety of applications without interference from other users, does not support maximum use of the limited spectrum available for military operations. To address this, a number of alternative approaches have been proposed that provide varying degrees of oversight from the spectrum managers. Some of these, including dynamic spectrum access managed by a central authority and opportunistic spectrum reuse by autonomous radio networks, are based on the conventional interference-avoidance strategy. Many modern radio devices are, in fact, able to tolerate quite high levels of interference, so one option for effective spectrum access is to allow users to co-exist, with limited regulation, by adapting their radio parameters according to the interference and propagation conditions they encounter. In this paper, we discuss these options, and provide some insights about the potential challenges and advantages of each.

1.0 INTRODUCTION

In conventional spectrum management, users are assigned spectrum slots and must follow certain rules regarding geographical area, power emissions and so on. In the civilian context, assignments for specific types of applications are grouped within allocations determined by national regulators – these are also agreed within the International Telecommunications Union (ITU). The objective of the spectrum planning and management process is to provide interference-free communications to licensed spectrum users.

National defence forces typically have their own spectrum management capability, which must work in cooperation with the national regulator in the home country. When in operations in other nations, military forces must follow the regulations of the host nation, and must coordinate with their coalition partners to assign the spectrum slots. This is a centralised coordination process, which develops a static assignment plan to accommodate each nation's systems. This process becomes increasingly complex as operations become larger and more numerous and rely on more sophisticated technology. The problem of effective spectrum management is therefore of particular interest to an alliance such as NATO.

The continuing increase in demand for commercial wireless services puts pressure on the spectrum available for military use. National auctions often bring in billions of dollars as portions of the spectrum are leased for commercial use. The spectrum is therefore viewed as a valuable resource, and defence forces in many nations are losing their own spectrum allocations as they are leased to support national government budgets. At the same time, new military technology requires more spectrum access to support command and control, surveillance, weapon control and other functions.

In this context, there is a need both to increase the efficiency of spectrum use, discussed in the companion paper [1], and to reconsider the fixed assignment approach to spectrum management. This management strategy may leave portions of the spectrum unused in some geographic areas for some periods of time, while at the same time there is demand from other users for spectrum access. Furthermore, many modern radio communication systems can tolerate some degree of interference, therefore planning for interference-free operations may be wasteful of the scarce spectrum resource.

In 2002, the US Federal Communications Commission (FCC) issued a report addressing the growth in the use of unlicensed devices that operate in the so-called "Industrial, Scientific and Medical" (ISM) bands [2]. These devices cover a range of consumer services, including garage door openers, cordless phones, baby monitors and security alarm systems. In this report, the Unlicensed Devices and Experimental Licenses Working Group of the Spectrum Policy Task Force recommended that unlicensed devices be permitted to make opportunistic use of spectrum licensed to existing services.

In 2003, the US DoD DARPA started a program in neXt Generation (XG) communications to develop *opportunistic spectrum access* technology. The aim was to take advantage of licensed spectrum that was unused, perhaps temporarily, by allowing devices to sense the available spectrum and determine whether the licensed user was present. These opportunistic devices would be allowed to operate, providing the interference they caused to the licensed user was kept within specified limits.

The IEEE set up the P1900 Standards Committee in 2005 to develop standards for radio and dynamic spectrum management. This was reorganised as the Standards Coordinating Committee 41 for *Dynamic Spectrum Access Networks* in 2007.

These developments have spurred a new direction in communications R&D. There has been an avalanche of papers in the academic literature about "cognitive radio" and dynamic spectrum access over the last few years. The term "cognitive radio" was coined by Mitola [3], in the sense of a truly intelligent, self-aware capability for software radio devices, integrating control, learning and language ability. The term has since been used and abused so much that it has lost a common meaning, but is widely thought of as meaning a device capable of exploiting spectrum opportunities.

After 10 years of investigation, there are even more questions about the potential and implementation challenges of dynamic spectrum access strategies. Ideally, the goal would be to transition from today's approach, which exacerbates a shortage of spectrum by enforcing an inflexibility in spectrum access, to a dynamic, effective approach that will allow adaptive planning for tactical operations, reduce spectrum fratricide and provide assured access to spectrum on demand. This is an important topic in many nations, and the NATO IST board has assigned Research Task Groups (IST-077/RTG-077 completed its work in December 2011, and IST-104/RTG-050 starts in 2012) to consider the role of cognitive radio in NATO [4].

In this paper, we consider some issues related to future spectrum management strategies for military use. We look first at dynamic centralised spectrum management, in which spectrum assignments are opportunistic but controlled by a spectrum management authority. Then we consider more distributed approaches to dynamic spectrum access, in which the decisions are made remotely by the radio devices themselves. The first case we look at is that in which opportunistic users are allowed to access licensed spectrum on a non-interference basis. Then we consider the situation where users co-exist without constraints on interference, either non-cooperatively or cooperatively, which requires less oversight from a spectrum manager.

2.0 SPECTRUM ACCESS STRATEGIES

It is unlikely that the current spectrum management pillar of spectrum assignments could ever be abandoned completely. A feasible aim would be to make the spectrum access more effective overall through the introduction of dynamism, and to increase this dynamism as the technology develops. This means that dynamic and fixed spectrum access users will need to co-exist, and this could be achieved either within the same spectrum or through a division of the spectrum to support different types of users. The latter approach is more consistent with current spectrum management philosophy. In either case, it is expected that, for the foreseeable future, most users would continue to be assigned spectrum slots, but that these might be chosen to provide their minimum bandwidth requirements. Users requiring additional spectrum to support higher bandwidth applications would then have to make use of spectrum opportunities, either by finding unused assigned spectrum or by accessing a shared pool of spectrum.

In one approach to dynamic spectrum access, the whole available spectrum continues to be assigned to licensed users (often called *primary users*), as is done currently in most spectrum bands, and other users (*secondary users*) are allowed to use vacant spectrum opportunistically on the condition that the interference they cause the license holder does not exceed specified limits. When the primary user has need of the spectrum, the secondary user must cease use within a specified time. This means that the secondary user must have good power control to satisfy the interference restrictions, must be aware of the presence of the primary user, and must also be able to adapt its operating frequency or spectrum occupancy to be able to move to another portion of the spectrum or to adjust its spectral mask to avoid the primary signal. The secondary user may be made aware of the primary user either through a centralised controller (Section 3.0) or through its own local sensing capability (Section 4.0). This reuse of licensed spectrum assignments is the strategy specified by the FCC in the broadcast TV bands, as defined by IEEE 802.22.

The second approach, the shared pool concept, is that spectrum is not assigned to specific users, but that all users access it according to certain policies, which may constrain emission power and bandwidth, for example, and interference levels are not guaranteed. This is like the ISM bands currently allocated within the UHF and low SHF bands.

For military applications, the primary/secondary approach would seem to be the most straightforward strategy for the introduction of dynamic spectrum access. Much of the spectrum is assigned to users that have limited flexibility in their spectrum access and, of course, these legacy users do not have the capability to sense the spectrum to identify alternative opportunities. This spectrum re-use is the strategy pursued within the DARPA XG program, in which secondary users were to sense the spectrum and make their own decisions about spectrum opportunities.

However, one of the significant challenges of this approach in a military environment is that spectrum is itself a battlefield. A secondary user is vulnerable to spoofing, in which an adversary emits signals that mimic the primary user, either to prevent secondary use of the spectrum opportunity in the first place, or to force the secondary user to move by making it appear as though the primary user turns on [5]. An intelligent adversary could cause significant disruption to the secondary user in this way.

Furthermore, this may be a short-sighted approach to introducing dynamism into spectrum access, and may limit effective access in the future. Specifically, it considers only the interaction of primary and secondary users. As dynamic spectrum access technology matures, and is more widely accepted, the challenge will be the interaction of unlicensed users, i.e., secondary and secondary. It may be determined in the future that these secondary users should co-exist by adapting their waveforms and power (Section 6.1), or they should co-operate (Section 6.2), or that some other access policy should be applied. However, the first generation of dynamic spectrum access devices will be the legacy users of the future, and careful thought should be given to ensuring that they have sufficient capabilities before their initial introduction, so they do not hamper the introduction of more sophisticated technology later.

The ISM-like band is particularly attractive for military applications. Concerns about protecting legacy users would be alleviated by preventing them from being assigned in this band. Users would not have to immediately vacate their spectrum when interference is detected, because the other user has no more rights to the spectrum. Modern digital waveforms can be designed for high interference tolerance, and users could determine their own thresholds for acceptable spectrum quality based on their applications' requirements. This would increase the robustness of communications, reducing the overhead of each user as it searches for more suitable spectrum and is forced to re-establish connectivity at a different frequency. This overhead reduction is a factor in more efficient use of spectrum, which provides more usable throughput overall.

The experience of the commercial ISM band is that its existence has precipitated the rapid development and introduction of many new technologies. A military ISM-like band would probably similarly encourage the development and implementation of effective dynamic spectrum access strategies, because uncertainties about regulations and market potential would be reduced. The technology would necessarily support multiple dynamic spectrum users, and would therefore skip the first generation problem identified above for the primary/secondary scenario.

3.0 CENTRALISED SPECTRUM ACCESS MANAGEMENT

The *radio environment map* (REM) concept is intended to maintain centralised management of spectrum access while introducing controlled, dynamic access to spectrum that is unused in certain geographic areas. This centralised control has advantages in providing the ability to protect certain users, for example, those in emissions control that would not be detected by spectrum sensing alone. The main disadvantage, however, is its complexity and difficulty in maintaining a current picture of the radio environment. The

REM is essentially a database, perhaps presented in a visual map-like representation, of users and their locations and radio parameters such as frequency, bandwidth, power and possibly modulation. We consider two approaches to generating the REM and their potential for application in a multi-national military context.

3.1 Spectrum registration

In a registration-based REM, spectrum users must submit their locations and radio parameters to the central authority, and this information is listed in the database. Users requiring spectrum access must submit their requests to the central authority, which can determine appropriate frequencies and operating parameters for the specified locations, and can then issue temporary leases. This process might be partially or fully automated. This is essentially the solution that has been determined within IEEE 802.22 for the reuse of TV broadcast bands. The protection of the licensed TV broadcasters parallels the need to protect legacy military users from interference.

Even though, as with the current spectrum management approach, not all users will abide by spectrum regulations, this is a promising strategy for the introduction of dynamism into spectrum access without giving up centralised control. One of the challenges in a military context is that this approach is most effective if it is shared among all the participating nations, but not all users will want to register all their spectrum users for other nations to see. Different levels of information sharing and security might have to be incorporated into this model. Another challenge is that this approach does not address the complexity of assigning the spectrum to support large operations and many spectrum users; in fact, it becomes more complex as this must now be done in real-time.

When the emitters' locations are established, a REM can be generated using a propagation prediction program, such as CRC-COVLAB [6], which uses information about the terrain and ground covering to predict the power levels at any location. This can be computationally intensive, but can be fully automated.

3.2 Sensing

Another approach to generating a REM is to use a network of spectrum sensors that measure the signal level across the frequency band of interest and forward that information to the central authority. Processing of the measurements then provides a spectrum occupancy map showing the power levels at each location. There has been a great deal written about the sensed REM approach, but very little experimental work to assess its feasibility.

There are several challenges with using sensing to generate a REM. The accuracy of each sensor measurement will impact the overall REM reliability: this accuracy can be affected by the local environment as well as by the characteristics of the emissions. To obtain a reliable REM, sensed information will be required from many locations within the geographical area of interest; this information may be received from mobile sensors or from many static sensors, but in either case, there is a large volume of information that must be passed to the central authority. In many military contexts of interest, the emitters will be mobile, in which case the sensor information must be updated sufficiently often to maintain a reliable REM. When all the sensor levels have been collected, it still remains to produce a REM from the information.

3.2.1 Propagation environment

A static sensor can measure the power level only at a single position, and this can be affected by multipath fading as well as shadowing. As an example of power levels measured over a small area, see Figure 1. In this case, the receiver was moved over a grid of 10×20 m, at 1 m intervals, and a transmitter operating at 370 MHz was located in a 2nd storey window approximately 200 m away. There is a very large variation, approximately 18 dB, in measured signal power over the receiver grid. There is a distinct boundary between the power levels in the top 5 m compared to the lower 15 m – this is attributed to shadowing. The variations within these areas are attributed to multipath fading – in more complex propagation environments, such as urban areas, these variations can be 30 dB or more over a single wavelength.

This illustrates the challenges with sensing power levels with a static sensor, which cannot average over a local area. A sensor mounted on a mobile platform reduces the very localised effects, but fairly sophisticated signal processing is required to identify when substantial changes occur, such as at the shadowing boundary shown in Figure 1. When multiple sensors are used, even when they are located some distance apart, the processing of the collected data can help to remove some of these discrepancies [7], but the resulting power level estimate will always be subject to uncertainty. The scale of this uncertainty determines the buffer that must be allowed in the spectrum access protocol; the trade-off is that over-estimating this buffer reduces the possibility to reuse spectrum, and thus reduces the overall spectrum access effectiveness.

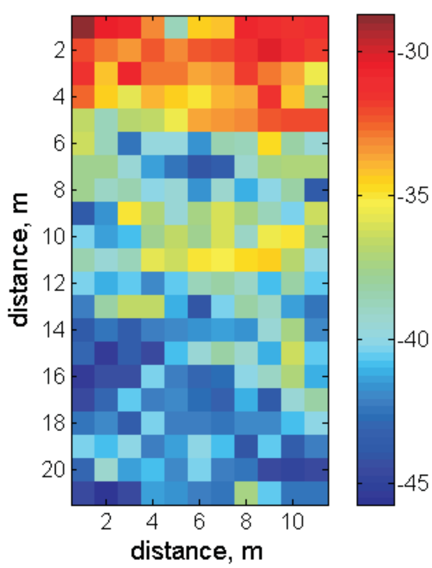
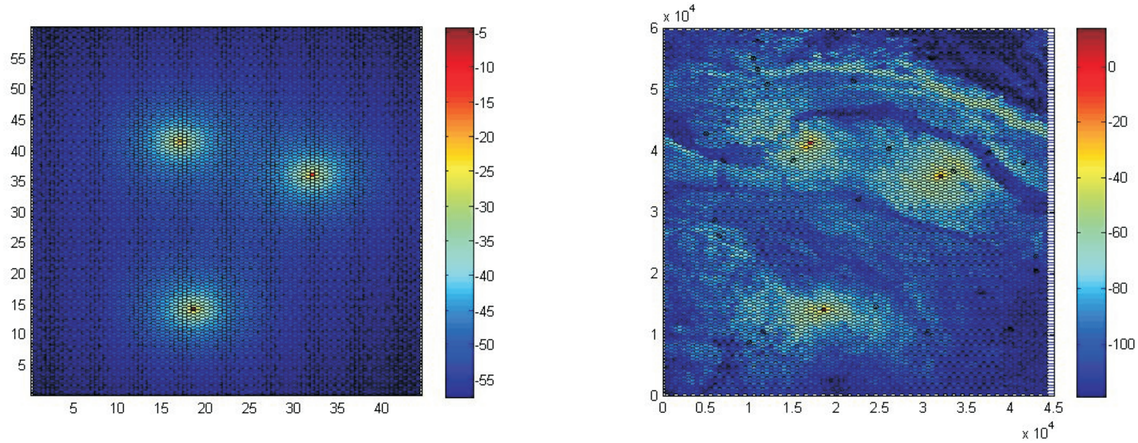


Figure 1: Variation in power measurements over an area of 10×20 m.

Figure 2 illustrates the challenges of real propagation environments. The RF power maps show an area of 45 km-by-60 km, with three transmitters operating at 300 MHz and 100 W. The left panel shows a "flat earth" model, in which no consideration is given to the effects of terrain on power levels. The right panel shows the power map for a real terrain model, located just west of Ottawa, where the terrain varies from about 100 m to 400 m above sea level. This was generated using CRC-COVLAB [6]: there is no way of knowing the accuracy of this map without extensive, and expensive, measurements on the ground.



(a) flat earth model (b) terrain-based model

Figure 2: **RF power maps for a** 45×60 **km**^{2} **area in eastern Ontario.**

Using the artificial REM of Figure 2(a) in place of the real REM in Figure 2(b) will cause unintended interference and will fail to achieve effective spectrum access. In more complex terrain, such as hills or mountains, or in built-up areas, the shadowing that complicates the real REM is even worse.

3.2.2 Emission characteristics

Most of the spectrum sensing R&D reported in the academic literature addresses the case where a primary, or licensed, user is either ON or OFF, with the assumption that a transmitter that is turned on has a 100% duty cycle, i.e., it is emitting power continuously. In the type of military, heterogeneous environment we are considering, this is unlikely to be the case. If we observe a single emitter which is part of a TDMA network, it might transmit in one slot out of every N. Unless our sensor is synchronised to that TDMA protocol, this means that the signal power will be underestimated by $10 \log_{10} N$ dB. If other users are also transmitting in that TDMA cycle, the sensor will receive signals of different power levels. A typical sensing objective is to determine whether the signal level exceeds a given threshold or not, hence the duty cycle effect can significantly impinge on the sensor's capability to achieve that goal.

Other real challenges that are typically avoided in the academic literature are the cases where the emitter location is not known or where there is more than one emitter (in this case, we may be able to simplify the problem by considering a network of radios in a confined geographic area to be a single emitter, and are then concerned with multiple networks within sensing range). The concept of primary emitter has been derived from the reuse of TV bands, in which case not only the precise location of the transmitters, but also their transmission power and signal characteristics are known. In the mobile, heterogeneous environment expected in military operations, none of these pieces of information may be available.

The problem of multiple emitters is particularly important where we are accepting that users will experience some interference, and therefore there may be several radios operating in the same bandwidth. This challenge is exacerbated because the problem is usually formulated as trying to identify the emitters' locations. In fact, we are not particularly interested in the location of the radios themselves; rather, we are trying to avoid interfering with potential receivers, therefore we really want to know where the signal power exceeds a prescribed threshold.

3.2.3 Collecting sensed information

When a network of spectrum sensors collects information over a large geographic area, that information must be sent to the processing centre for processing and storage. Simulations of this process using different routing protocols and optimistic assumptions about propagation characteristics show that there can be substantial delays in receiving information, as packets may be lost due to congestion or require many hops to arrive at the destination. This gets worse as the density of sensors increases, but if the number of sensors is too small, the quality of the resulting REM is poor [8].

If the sensors are mobile, fewer are required to cover the area, but their sensed information must still be forwarded in a timely manner to the processing centre, thus there is a lower bound on the amount of bandwidth required. If the emitters are also moving, then the time window in which the sensed information must be received is even shorter [9].

3.2.4 REM generation

The question of generating a REM from sensed data is still not adequately answered. Techniques that appear promising for the flat-earth model, Figure 2(a), do not necessarily work for the more realistic models such as that illustrated in Figure 2(b). The REM generation is particularly difficult when there is no additional information about the emitters, such as location and power. Some information might be available from other domains, such as a land or air common operating picture, however these do not give a complete picture of operations so an open question is how to combine sensor input effectively with supporting information.

Another approach to create a REM, which avoids many of the limitations outlined above, would be to use airborne or satellite sensing, which could provide good estimates of both location and power. The REM could then be generated using a propagation prediction program such as CRC-COVLAB [6], which uses readily available terrain maps. This, however, requires costly infrastructure.

4.0 DISTRIBUTED OPPORTUNISTIC SPECTRUM ACCESS

An alternative to having a central authority retain responsibility for assigning spectrum is to delegate the selection of spectrum opportunities to individual users and networks. Again, we are interested in a heterogeneous, mobile environment. We consider in this section the case where the spectrum slots have been assigned as in conventional spectrum management, but opportunistic users are allowed to access the spectrum if it is unused.

We assume that the behaviour of the opportunistic users is constrained in some way by rules, or policies, defined by a central authority. These might include restrictions on the spectrum slots that may be used – this can provide protection to particular users such as radar, or to users that are in emissions control mode and therefore rarely emit detectable power. Policies might also cover power detection thresholds, number of slots accessed simultaneously, duration of opportunistic access, etc.

We consider the case in which a network of opportunistic radio nodes is seeking available spectrum slots, and has been informed by some form of policy that the radios must give way to another user in those spectrum slots. The opportunistic nodes do not know the location, power or characteristics of the signal, or even the number of emitters within the spectrum slot. This situation is unlike the TV band case (Section 2.0), where there is a well-defined primary user. To avoid the rendezvous problem, in which radio nodes wander across the spectrum looking for their network colleagues [10], we also assume that the opportunistic users have access to another spectrum slot, possibly shared with other users, that can be used for information exchange. Such an approach does introduce vulnerabilities, because now a malicious user needs only to block access to this control channel, but this can be handled using various signalling strategies.

The strategy for combining the sensing information of multiple nodes will depend on their mobility. If the nodes are static, they have no ability to average over multipath fading, so more information is required from adjacent nodes to counter the fading nulls, which may be quite deep. For mobile nodes, averaging at each sensor can overcome the multipath fading effects, then the combining must counter only shadow fading, which is typically shallower and has higher spatial correlation.

4.1 Small networks

When there is a small number of nodes in the opportunistic radio network, they can work together to overcome the challenges of the propagation environment illustrated in Section 3.2.1. The typical approach to this is to apply an OR-type algorithm: the sensors apply a local threshold to determine whether detected signal power is significant and then a decision is made that the signal is present if any one (or some specified number) of the sensor power levels exceed that threshold. This is efficient from the point of view of the amount of information exchanged – each sensor can use a single bit to indicate presence or absence.

The performance of a dynamic spectrum access system is not only measured by the probability of detecting the presence of the licensed user, but also by the probability of a false alarm, i.e., the probability that the opportunistic network decides the licensed user is present, when in fact it is not. This is a costly mistake, as it results in the opportunistic network unnecessarily vacating the spectrum slot, losing bandwidth capability, and needing to find an alternative spectrum slot elsewhere. The OR algorithm is particularly bad for this: the presence of noise can make any sensor make a bad decision, and this affects the performance of the whole network sensing.

To average out the effects of the noise at each sensor, some kind of averaging across the sensor measurements is required. There is a variety of average consensus algorithms that achieve this goal, improving overall performance by reducing the false alarm threshold and improving the reliability of detecting the licensed user's signal. Although more information must be exchanged between nodes, it only needs to be transmitted amongst immediate neighbours to perform the averaging. This class of algorithm also averages out any variation in the power levels across the nodes, which may not be appropriate. If the radio nodes are spread over a large area, then shadowing effects may result in different power levels being detected at each node. The averaging then reduces the sensitivity of the processing algorithm to those nodes that measure a high level of the emitted signal. More environment-aware cooperative sensing algorithms are required that avoid this problem, an example is presented in [11].

One of the disadvantages of decentralised algorithms of this type is that they are iterative, meaning that information must be repeatedly updated and forwarded. The challenge is to achieve a robust decision in as few iterations as possible. Note, however, that the scenarios usually considered in work reported in the literature is of static emitters and opportunistic radio nodes. In the environments we are interested in, the emitters and nodes may be mobile and the emissions may be transient. In this case, the iterative behaviour of the algorithms can form part of the necessary tracking of changes with time.

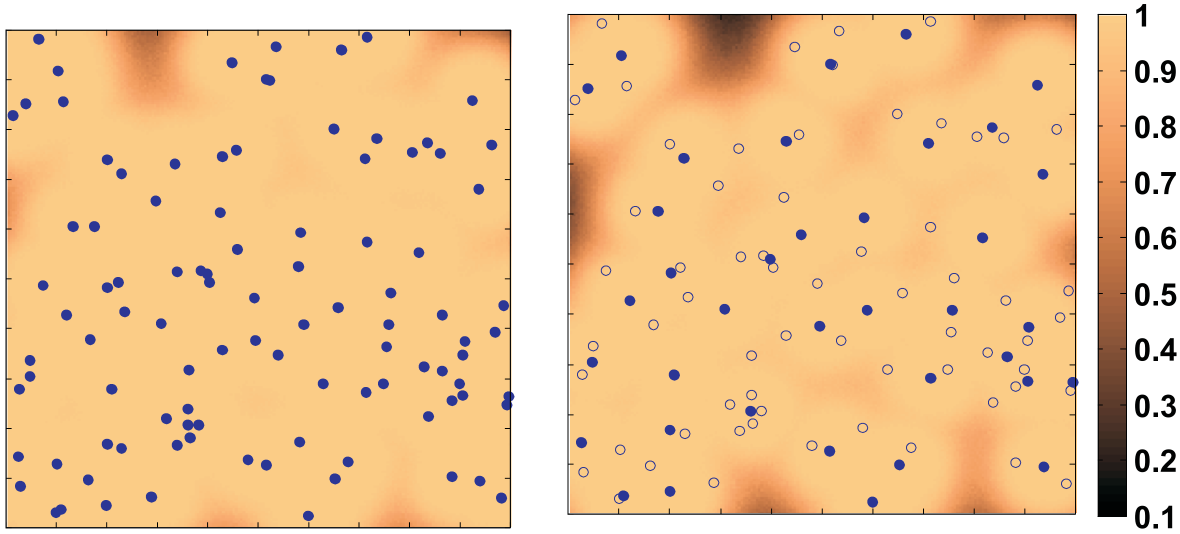
4.2 Large networks

The challenge for the opportunistic user is not only to detect whether or not the spectrum slot is occupied, but also to achieve this in an efficient manner. Any exchange of information among the radio nodes requires use of a control channel, which may be congested or bandwidth constrained, and the need for which impacts the overall effectiveness of the spectrum use. This is particularly a problem for large networks, where multiple hops may be required to pass information from one node to another. Some nodes in a local area will have similar power level measurements, hence the limited benefit gained by receiving observations from all of them will be more than offset by the cost of sharing that information.

To reduce the number of sensing radio nodes that must share their spectrum observations, the observing nodes can be clustered into groups that share similar observation characteristics. One representative of each group then interacts with other such representatives, to form a small network as in Section 4.1. One method to do this clustering is the K-means algorithm [12], which groups the nodes by their locations; this grouping is done at a central location, meaning that all information is passed to a central node for processing and then the clustering decisions are fed back to the remote radio nodes. As observed in Section 3.2.1, sensors that are located close to one another may not share the same shadowing profile, therefore grouping solely by geography may not provide adequate coverage.

A distributed approach to clustering was introduced in [13] to address these challenges. This uses a message-passing algorithm in which the nodes share their information only among immediate neighbours and make local decisions regarding cluster membership based on similarities in received power. The success of this algorithm is illustrated in Figure 3, where the probability of detecting an emitter in any location is indicated by its brightness (the parameters for this simulation can be found in [13]). The left panel shows the results for including all nodes, while in the right panel, 37 of the total 100 nodes have been selected using the message-passing algorithm. These results ignore the effects of multipath fading, and consider only shadowing: this would be applicable when the nodes have even a small amount of mobility. It is clear that the well-selected nodes provide a coverage that is almost as good as when all nodes are used, but the amount of information that must be exchanged beyond the one-hop neighbourhood is drastically reduced.

Even though this algorithm is distributed, meaning that the decisions about clustering are made locally rather than centrally, this algorithm provides results that are as good as, or better than, the K-means algorithm. Unlike the K-means algorithm, the location of each sensing node is not required, and the nodes are not required to communicate with a central processing node: both of these features are particularly useful when the sensors are mobile.



(a) All nodes (b) AP-clustered nodes

Figure 3: Probability of detection, shown by brightness, of licensed user achieved by reporting nodes (solid) selected from amongst all nodes (circles).

5.0 EVALUATION

It is not clear that allowing opportunistic access will necessarily improve the result of spectrum access, and it is to be expected that different spectrum access concepts will yield different behaviours. Early attempts demonstrated only the capability of a single network to move around other spectrum users without causing interference [14]. However, in an environment with more than one network of opportunistic users, it is necessary to measure the abilities of all networks to provide the required quality of service simultaneously, as well as to evaluate higher level qualities such as robustness, resilience, and efficiency [15].

Because effective spectrum management is not concerned with a single network of radios, but rather all the users that may occupy the spectrum band of interest in a given geographic area, it is important to consider the effects of dynamic spectrum access on the whole community of interest.

Simulating different approaches, assessing the resulting performance and measuring the overall effectiveness is not a simple problem. To evaluate each radio and network fully and separately would require a massive amount of simulation and complex analyses. The NATO IST Research Task Group (IST-077) has recently proposed an approach using a scenario-based evaluation tool [4]. The vignette includes multiple national participants, platforms and applications, as illustrated in Figure 4. Rather than evaluating all the radio and network parameters of each user, the simulation method increasingly stresses the spectrum by successively introducing new requirements, and each is measured simply as successful or not, based on generalised metrics. This allows a relatively straightforward comparison of spectrum access protocols and will facilitate identifying particularly challenging operating conditions. This work is continuing within the subsequent working group, IST-104.

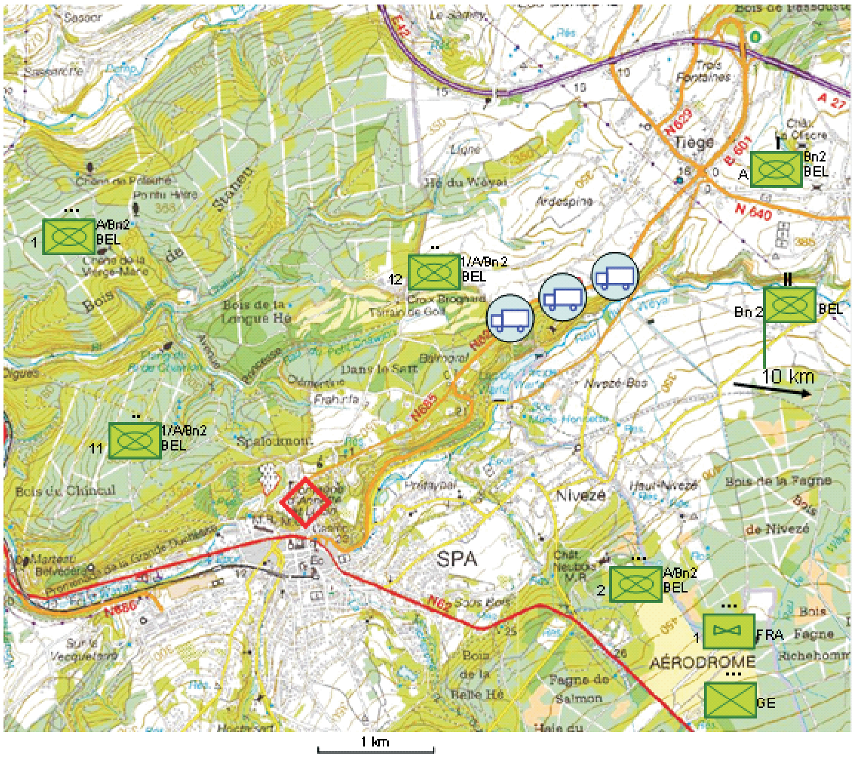


Figure 4: **Vignette used in simulating and evaluating dynamic spectrum access concepts [4].**

As the condition for accessing unused spectrum slots is that opportunistic users vacate them when the licensed user appears, the nodes must continue to sense while they are operating, and must be able to distinguish the signals of their fellow network nodes from the licensed user. Further, without knowing the signal characteristics of all possible licensed emitters, sensors must employ energy detection. However, energy detectors are unable to distinguish between different types of emissions. It was seen in experimental measurements reported in [16] that these detectors will be unable to separate out-of-band emissions and other noise sources from user signals. Malicious spectrum users can also cause them to vacate simply by emitting signal energy within detection range. Opportunistic users in licensed spectrum will therefore be vulnerable to abandoning spectrum slots unnecessarily, and in doing so will suffer unreliability, unpredictability, and additional overhead costs to find a vacant spectrum slot, perform network connection operations, and maintain vigilance for the return of the licensed user. There are additional security issues related to the use of dynamic spectrum access, see [17] for some examples.

The so-called *hidden node* problem is fundamental for dynamic spectrum access. Although the nodes sense the transmitted signal, they have no information about the location of the intended receiver of that transmission. Any transmission from the opportunistic user will interfere with the receiver, not the transmitter, so it is necessary to leave a buffer in the power detection threshold to account for this interference.

While policies can, as noted above, be used to restrict the behaviour of the opportunistic users, these must be monitored and enforced. It is particularly challenging to monitor whether a sensing node, or network of nodes, is properly applying a power detection threshold, as the detected levels are highly specific to the locations and algorithms applied.

6.0 SPECTRAL OVERLOADING

So far, we have considered the objective of spectrum management to be the assignment or opportunistic use of channels to provide interference-free, or nearly interference-free, radio operation. In fact, many signalling schemes are quite tolerant of interference, especially when implemented with advanced signal processing at the receiver. This leads to the concept of a military ISM-like band, in which opportunistic access is encouraged, but no guarantees about spectrum quality are provided, and users can apply spectrum access protocols of their own choosing.

The objective of the sensing and adaptive radio is then to provide channels of sufficient quality only to provide the desired quality of service, for example, in terms of data and error rates. In this case, we can conceive of allowing users to cause quite significant interference on the basis that the interfered users can adapt their parameters as necessary to accommodate it.

We refer to this approach as spectral overloading, and consider two examples here. The first is arrived at using a resource allocation algorithm, in which both users autonomously adapt their waveforms, powers and data rates to maintain connectivity, without explicitly cooperating with each other. In the second, the users rely on a distributed cooperation and exploit the multipath propagation to design their own waveforms to control the interference all users experience.

6.1 Non-cooperating users

Resource allocation problems are typically analysed in the academic literature using a game theoretic approach based on the Shannon channel capacity. However, this gives a warped evaluation of how two pairs of users can co-exist in the same spectrum slot as the assumptions about the signal characteristics required for the analysis are not met in practice. An algorithm was developed in [18] that uses real signal characteristics, in which two user pairs react independently to changes in the interference levels they cause each other, each aiming to maximise its data rate while maintaining a specified bit error rate. This algorithm operates iteratively, so that the two user pairs converge to a selection of power, waveform and data rate that meets their required error performance. While an iterative approach might appear to be complex and time-consuming, as with the distributed sensing algorithms discussed in Section 4.1, it is actually well suited to continuously updating the radio parameters when the users are mobile.

Figure 5 shows the radio parameter selections for two user pairs at various separations, taken from [18]. Each is allowed to transmit with a maximum of 5 W, and has the choice of BPSK-OFDM, QPSK-OFDM and direct sequence spread spectrum (DSSS) using BPSK modulation with spreading sequences of length 15 (SF-15). The transmitter and receiver of each pair are assumed to maintain a constant separation. The channel model is Rayleigh fading using a standard pathloss model at a frequency of 2 GHz.

When the user pairs are 1 km apart, the interference is quite low, and each can operate using OFDM with QPSK to achieve the highest possible data rate. Neither user transmits at full power: the 'circle' user uses less power simply because of an artifact in the algorithm, that the last user in the iterative updating has a better knowledge of the interference characteristics than the first. This would not appear in a system that updates continuously in a mobile environment.

As the user pairs get closer together, they have to increase their power to counter the effects of interference from the other transmitter to sustain the desired error rate. At a separation of around 620 m, the circle user (again, because it makes the last selection) switches to BPSK, and as a result, needs much less transmit power. Soon after the separation decreases below 500 m, the circle user switches to DSSS-SF15. This requires a very low power to achieve the desired error rate, hence the interference caused to the 'square' user is reduced, which allows it to reduce its power while retaining the maximum data rate. When the separation is just over 100 m, the square user switches to BPSK, and finally, to DSSS-SF15.

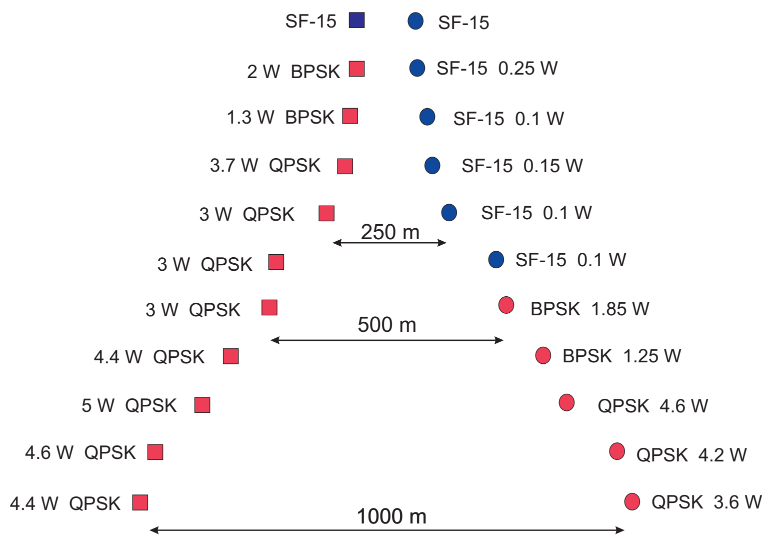


Figure 5: **Power, waveform and data rate adaptation for spectrum co-existence of two user pairs (red = OFDM, blue = DSSS).**

This example illustrates that the two user pairs can co-exist, and maintain an acceptable error rate even though they cause each other considerable interference. Even though the data rate may be very low, connectivity is maintained, which reduces or eliminates the cost (power, time, bandwidth) of finding another spectrum opportunity and reconnecting there. In a mobile environment, this is a practical strategy – the interference is likely to be fleeting, hence tolerating the interference is a more effective approach to spectrum access than avoiding it.

6.2 Cooperating users

Interference alignment (IA) is a technique that exploits spatial diversity in multipath channels, and is related to MIMO, discussed in the companion paper [1]. In IA, the capability of multiple antennas at the transmitter and receiver is used to support multiple links between pairs of nodes simultaneously. The original concept [19] was to use cooperation at the K transmitters, each with N antennas, to select spatial signatures such that each receiver, also with N antennas, is able to remove all the interference using spatial processing. This is achieved, in theory, by separating the signals in the spatial domain.

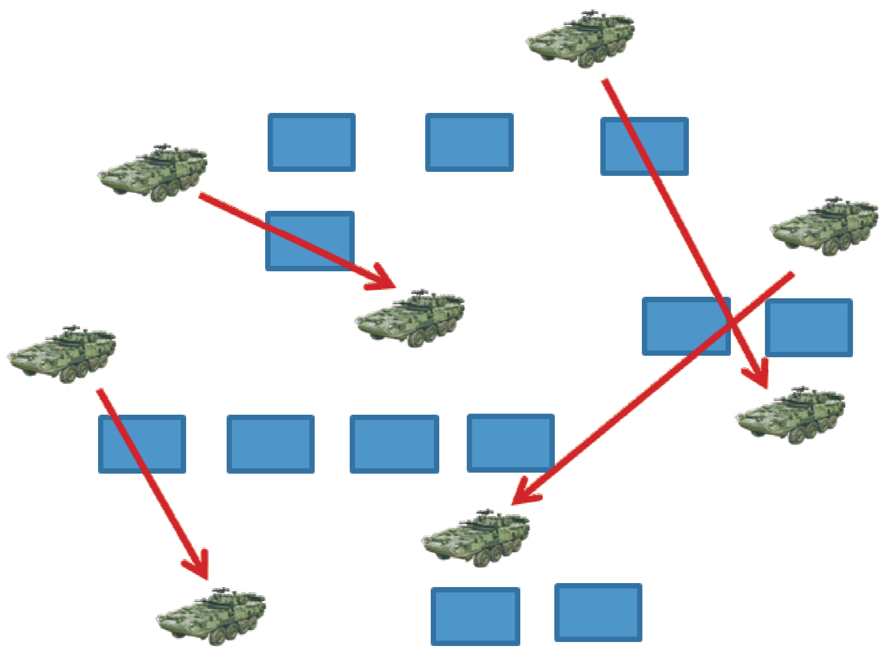


Figure 6: SINR distributions resulting from spectrum overloading using interference alignment.

This spectral overloading technique can, in theory, provide $K = 2N - 1$ pairs of users with interference-free links, where each user has N antennas for transmitting or receiving. Contrast this with adaptive antenna arrays at the receivers only, which are able to null out up to $N - 1$ interfering signals.

This ability to overload the spectrum might seem magical. With enough antennas at each node, huge numbers of links could be supported within a small local area in a single channel bandwidth. There is a catch, of course. The theory requires that all transmitters have up-to-date information about all the links. For a narrowband (frequency-flat) channel with K pairs of users, each of the K receivers must share the $N \times N$ complex channel gain values obtained from each of the K transmitters. Furthermore, for a mobile environment, these K^2N^2 complex values must be shared often enough to remain timely.

In recent work [20], [21] on spectral overloading using interference alignment, algorithms have been proposed that are based on more realistic assumptions about the communication system and the operating environment. In particular, the amount of information exchanged by the receivers is reduced to a feasible amount. The aim is not to eliminate interference, but rather to limit it to a manageable level.

The scenario is illustrated in Figure 6. Several pairs of users are communicating at the same time, in the same bandwidth and the same general location: a scenario in which interference and multipath are significant problems to conventional radios. In the proposed method, each receiver estimates the $N \times N$ complex channel matrix response from each of the K transmitters. Each receiver instructs its transmitter which precode it prefers, as in [1, Section 3.1]. Other receivers in the neighbourhood also hear this information, therefore when all receivers have fed back their choices, all transmitters know the preferred precode for their own receivers, and all receivers know which precodes all transmitters will use. The receivers can then compute their expected performance for a small set of similar precodes and each feeds back a single bit for each to indicate if it would improve (+1) or degrade (-1) their performance. Each transmitter then selects the precode that has the largest sum of feedback scores, and after a small number of iterations, the receivers are each able to receive their intended signals with a tolerable level of interference. This is a distributed cooperation approach which minimises the amount of information fed back.

In a mobile environment, the channel responses change over time, hence this update process continues periodically to ensure that the interference levels are controlled for all receivers. As discussed in [1, Section 5.2], the rate of update is dependent on the rate of change of the channel matrix, which depends primarily on the environment, the speed and the frequency.

Simulations have been run for an urban non-line-of-sight environment at 2 GHz, in which $K = 6$ mobile user pairs were randomly placed in an area of $500 \times 500\ m^2$, where each user had $N = 4$ antennas. Figure 7 shows the distribution of SINRs obtained in the simulation for three spectral overloading approaches. A high SNR was considered to allow us to focus on the impact of interference rather than noise. The proposed algorithm (blue) achieves an SINR less than 2 dB lower than the “optimal iterative IA” algorithm (red), which assumes an infinite amount of feedback is available to provide all $N \times N$ channel response estimates to all transmitters and receivers to compute an optimal solution. A delay of 2 ms in the information fed back was included in both cases. The efficient algorithm provides a gain of approximately 7 dB over the case in which each user transmits with a fixed precode, as in ODMP (see [1, Section 3.1]), which uses no knowledge of the channel.

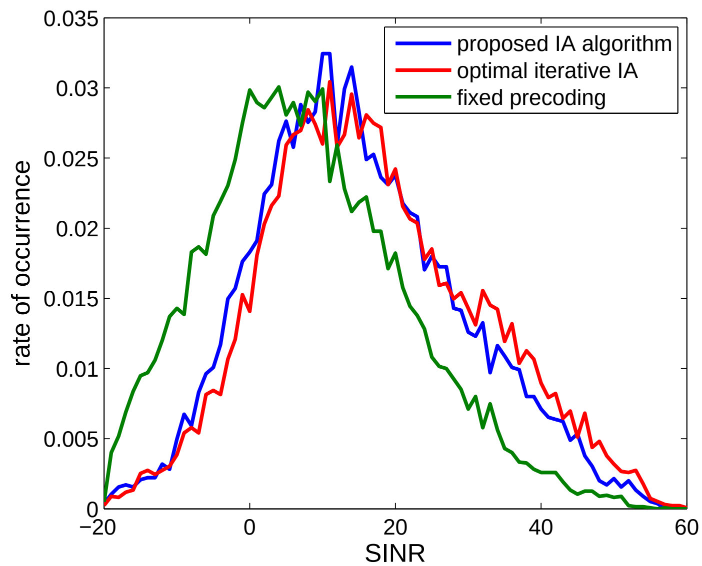


Figure 7: **SINR distributions resulting from spectrum overloading using interference alignment.**

The investigation of IA for spectral overloading has shown that, as might be expected, there is a reduced benefit from feedback as the channel changes more rapidly, i.e., at higher mobility. The errors in channel state information introduced by the mobility counter the gains provided by updating the precode selection. However, it is clear that there is potential to achieve much more effective use of spectrum through approaches such as cooperative interference mitigation, rather than assigning each user pair to a different spectrum slot.

7.0 CONCLUSIONS

As stated in the introduction, the aim is to achieve dynamic, effective approaches to spectrum management that will allow adaptive planning for tactical operations, reduce spectrum fratricide and provide assured access to spectrum on demand. The companion paper [1] addressed using the spectrum as efficiently as possible; smart information management is also important to reduce the demand on spectrum access.

Effective spectrum use is not a property of individual communication links, but considers the overall access to spectrum to support as many users as possible. Some increase in effectiveness is likely by allowing opportunistic users to access spectrum that is assigned (licensed) but is unused for a period of time. However, this is rife with challenges, including the need to have accurate and timely knowledge of the spectrum occupancy in the case of a centrally-managed system, and the reliability of sensed power levels in the distributed case. In both cases, the propagation environment introduces additional uncertainties which must be taken into account. Spectrum access strategies for this opportunistic reuse require tolerance to unpredictable and unreliable bandwidth for the opportunistic users, increased vulnerability to malicious and unintended interference, and the additional costs incurred for real-time spectrum management and/or exchange of additional information.

An alternative strategy is to dedicate some portion of the spectrum for dynamic access, in which no guarantees regarding the levels of interference are provided. This is the approach taken in the current ISM spectrum bands, where it has been seen that the ready access to unlicensed bands has encouraged the development of new technologies and applications. In these bands, users adapt their own radio and network parameters, cooperatively or non-cooperatively, to maintain their desired quality of service. There is a degree of unreliability to this approach as well, of course, as too many users in too small an area may result in unsustainable levels of interference to all.

There will remain a need for managed spectrum for users that need protection from interference. It is hoped that, as a result of future developments in radio and network technologies, the proportion of these users will reduce, and that more effective spectrum access will result. If not, the impact of spectrum congestion for military users will worsen, particularly in coalition operations.

ACKNOWLEDGEMENTS

Thanks to my colleagues K. Baddour, O. Üreten, G. Colman at CRC, and Maj. D. Jacob-David, for allowing their work to be presented here, and for supplying figures and other supporting information.

REFERENCES

[1] T. J. Willink, "Efficient access to radio spectrum," in *Proc. NATO RTO IST-105 Lecture Series on Next Generation Communications*, Apr. 2012.

[2] FCC: Spectrum Policy Task Force, “Report of Spectrum Efficiency Working Group,” Nov. 2002. ET Docket No. 02-135.

[3] J. Mitola, *Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio*. PhD thesis, Royal Institute of Technology (KTH), Sweden, 2000.

[4] C. L. Martret, “Cognitive radio in NATO: final report of task group IST-077/RTG-035,” Feb. 2012.

[5] Q. Peng, P. C. Cosman, and L. B. Milstein, “Spoofing or jamming: Performance analysis of a tactical cognitive radio adversary,” vol. 29, pp. 903–911, Apr. 2011.

[6] “CRC-COVLAB® V 3.0 coverage prediction and analysis software.” Available at: http://www.crc.gc.ca/en/html/covlab/home/home.

[7] M. Wellens, J. Riihijarvi, M. Gordziel, and P. Mahonen, “Evaluation of cooperative spectrum sensing based on large scale measurements,” in *Proc. DySPAN*, 2008.

[8] S. Faint, O. Ureten, and T. Willink, “Impact of the number of sensors on the network cost and accuracy of the radio environment map,” in *Proc. Can. Conf. Elect. Comp. Eng*, May 2010.

[9] S. Faint, “Effects of node mobility on radio environment map accuracy and network connectivity,” Tech. Rep. TM 2011-189, Dec. 2011.

[10] N. C. Theis, R. W. Thomas, and L. A. DaSilva, “Rendezvous for cognitive radios,” vol. 10, pp. 216–227, Mar. 2010.

[11] S. Barbarossa, G. Scutari, and T. Battisti, “Cooperative sensing for cognitive radio using decentralized projection algorithms,” in *Proc. SPAWC 09*, pp. 116–120, 2009.

[12] A. Malady and C. da Silva, “Clustering methods for distributed spectrum sensing in cognitive radio systems,” in *Proc. MilCom*, 2008.

[13] O. Ureten, K. E. Baddour, and T. J. Willink, “Distributed selection of sensing nodes in cognitive radio networks,” in *Proc. ISWCS*, pp. 1056–1060, 2010.

[14] M. McHenry, E. Livsics, T. Nguyen, and N. Majumdar, “XG dynamic spectrum access field test results,” *IEEE Commun. Mag.*, pp. 51–57, June 2007.

[15] T. J. Willink and H. Rutagemwa, “Framework for performance evaluation of cognitive radio networks in heterogeneous environments,” in *Can. Conf. Elec. Comp. Eng.*, May 2009.

[16] M. Wellens, J. Wu, and P. Mahonen, “Evaluation of spectrum occupancy in indoor and outdoor scenario in the context of cognitive radio,” in *Proc. CrownCom*, pp. 420–427, 2007.

[17] T. Newman, T. Clancy, M. McHenry, and J. Reed, “Case study: Security analysis of a dynamic spectrum access radio system,” in *Proc. GlobeCom*, Dec. 2010.

[18] D. D. Jacob-David, “Effective resource allocation for non-cooperative spectrum sharing,” Master’s thesis, Univ. Ottawa, Canada, 2011.

[19] V. Cadambe and S. Jafar, “Interference alignment and degrees of freedom of the k-user interference channel,” *IEEE Trans. Inform. Theory*, vol. 54.

[20] G. W. K. Colman and T. J. Willink, “Distributed MIMO interference alignment in practical wireless systems,” in *URSI Gen. Assembly*, Aug. 2011.

[21] G. W. Colman, "Distributed MIMO interference alignment in practical wireless systems." EU-COST IC1004, TD(11)02024, Oct. 2011.

ABBREVIATIONS AND ACRONYMS

BPSK	binary phase shift keying
CRC	Communications Research Centre Canada
DARPA	US DoD Defense Advanced Research Projects Agency
DoD	US Department of Defense
DSSS	direct spread spectrum sequence
FCC	US Federal Communications Commission
IA	interference alignment
ITU	International Telecommunications Union
ISM	Industrial, scientific and medical
MIMO	multiple-input, multiple-output
ODMP	orthonormal diversity multiplexing precoding
OFDM	orthogonal frequency division multiplexing
QPSK	quaternary phase shift keying
REM	radio environment map
RF	radio frequency
SHF	super high frequency (3 GHz – 30 GHz)
SIR	signal-to-interference ratio
SINR	signal-to-interference-plus-noise ratio
SNR	signal-to-noise ratio
TDMA	time division multiple access
UHF	ultra high frequency (300 MHz – 3 GHz)
VHF	very high frequency (30 MHz – 300 MHz)
XG	DARPA's neXt Generation program