Geolocation-Based Architecture for Heterogeneous Spectrum Usage in 5G

Oliver Holland, Member IEEE  
Centre for Telecommunications Research  
King’s College London  
London, UK  
oliver.holland@kcl.ac.uk

Mischa Dohler, Fellow IEEE  
Centre for Telecommunications Research  
King’s College London  
London, UK  
mischa.dohler@kcl.ac.uk

Abstract—It is becoming clear that one key characteristic of 5G communication systems will be heterogeneity. Heterogeneity is important to be able to reuse equipment that is available, extract maximum capacity from what is allowed to be used in given spectrum bands and locations due to regulatory limitations, and to match capabilities to the extremely wide range of use cases of 5G. Importantly, this heterogeneity will extend to the use of a range of different spectrum bands and spectrum types. Moreover, 5G must be supported extensively by lower-frequency options; millimeter-wave will be groundbreaking, but for propagation reasons cannot be the only solution to serve 5G, e.g., in achieving five-nines reliability. A key to realising lower-frequency spectrum for broadband (including mobile) communications systems is spectrum sharing. This has been emphasised by various reports by bodies such as the President’s Council of Advisors on Science and Technology in the US, and the European Commission in the EU. It is also noted that high-profile efforts such as TV White Space (TVWS) and the 3.5 GHz “Innovation Band” in the US imply geolocation database-assisted spectrum sharing as the solution to realising this extra capacity. This paper proposes the broad introduction of a geolocation database-based system to assist heterogeneous spectrum usage and spectrum sharing in 5G—both to enable 5G systems to use spectrum assigned to other services (e.g., TV spectrum, as in the TVWS example), and to enable 5G systems operating under different operators/owners to share spectrum with each other. This paper also argues for the geolocation databases to act as a management capability among secondary/opportunistic networks and devices, to better allocate resources, or combinations (aggregations) thereof, in the light of traffic requirements. The benefits of the approach are shown through reference to performances achieved by the authors within a major regulator-driven trial of TVWS in the UK. Further architectural observations are derived based on those performances.

It is important to note that this paper only proposes the high-level nature of such an architecture. Extensive detail is currently impossible to provide given that the basic architectural assumptions for 5G systems are not yet known.

Keywords—Spectrum databases, spectrum sharing, 5G.

I. INTRODUCTION

The basics of communication theory make it clear that vast bandwidths of additional spectrum are necessary to achieve the data-rates and other capabilities that are being proposed for 5G systems [1]. Millimeter-wave (mm-wave) technologies can realise vast amounts of new spectrum, and have some clear advantages such as the capability of very precise beamforming and likely scope for the introduction of a vast number of antenna elements in radios, leading to extremely high spatial frequency reuse and advanced MIMO techniques. However, it will always be difficult to maintain coverage using mm-wave technology, due to propagation issues. It will therefore be impossible for such high frequencies to realise the proposed five-nines reliability for 5G communication systems. Far lower-frequency spectrum is necessary for that.

In addition to regulatory efforts to make dedicated spectrum available for 5G systems, e.g., at the ITU World Radiocommunication Conference 2015, spectrum sharing is seen as a primary method to make more such lower-frequency “beach-front” spectrum available. This has been emphasized in a number of key reports (see, e.g., [2]). Moreover, to extract necessary capacity, heterogeneous spectrum usage will be necessary, comprising the combination of licensed, (conventional) unlicensed, (unconventional—e.g., white space) unlicensed, and light-licensed spectrum resources. 5G will additionally involve usage of a heterogeneous range of deployed systems, in order to take advantage of what is already out there, and in order to ensure that it is possible to use the associated types of spectrum if only given types of systems are allowed to be used in given spectrum, e.g., as is commonly the case for regulatory reasons. Such system and spectrum heterogeneity implies a large number of entities and devices interacting, having different owners and manufacturers, and that interaction has to be taken forward in an integrated/coordinated system taking into account traffic demands of 5G networks and devices such that the sharing can occur efficiently or occur at all.

Given the above observations, this paper introduces the concept of an integrated management architecture for heterogeneous spectrum usage (incorporating spectrum sharing) in 5G contexts. The purpose is to improve spectrum availability for 5G systems and enhance the optimality of their use of resources. We base this architecture on geolocation databases such as are being considered in the context of TV white spaces (TVWS) in various counties internationally, and in other contexts such as 3.5 GHz spectrum sharing in the US, and LSA spectrum sharing in Europe. This paper is structured as follows. Section II introduces the motivation for the geolocation database-based system, essentially based around a hierarchy of databases serving 5G systems in an integrated approach. Section III provides a suggestion for a broad outline of the supporting architecture—from a very high-level perspective noting that there are many uncertainties around 5G.
II. MOTIVATION FOR THE TECHNICAL SOLUTION

There have been extremely positive moves in the US, UK, Japan, Singapore, Kenya, South Africa, Tanzania, India, and elsewhere, towards opportunistic/localized sharing of white space, particularly TVWS (see, e.g., [3], [4]). These moves have underpinned the realisation that geolocation databases, in some cases assisted by local and standardised/reliable sensing mechanisms, can be practical solutions to unlock spectrum for opportunistic usage. Further moves are underlining such approaches. For example, the 3.5 GHz tiered spectrum access “innovation band” in the US is based on a “Spectrum Access System” (essentially, databases or similar cloud-implemented functionalities) supported by sensing [5]. And LSA approaches proposed in Europe, currently for spectrum sharing between mobile operators although eventually a lot wider than that, are key capabilities broadly also based on a database system that informs when/where the spectrum can be used by another operator, based on an agreement with the incumbent operator and a license (along with QoS guarantees) being awarded to the incoming/opportunistic operator [6]. Although this approach generally has far less dynamicity than others such as TVWS and in the 3.5 GHz innovation band, the licensing of the opportunistic operator and associated QoS guarantees for that operator are a beneficial trait.

Typically, database-based opportunistic secondary spectrum access systems do not resolve how the devices that are opportunistically using locally-unused spectrum will coexist, noting that there is significant potential for interference among these devices hence uncertainty in the quality of the spectrum they will see—particularly in TVWS. Such uncertainty must be resolved in order for some challenging 5G applications—such as Machine-to-Machine and the Internet of Things—to be realised with sufficient reliability of communication, and indeed to achieve the proposed “five-nines” reliability for 5G. Such a database system, in a 5G context, could be used to also manage the interference among the opportunistic devices, where it is noted that the UK/EU TVWS framework, for example, is already capable of that should it be required, only with changes to the calculations implemented in databases being necessary.

In addition to the above, it is noted that usage of unlicensed spectrum in general is typically very disorganised, and far better efficiency could be extracted from it through spectrum usage coordination among the unlicensed devices. Currently, mostly by chance of market dominance by a certain family of standards [7], the situation is less severe in 2.4 GHz ISM and 5GHz U-NII bands through the common use of Carrier-Sense Multiple Access (CSMA) interference avoidance schemes. Coordination of access among the 5G systems and devices that are accessing general (i.e., non-TVWS) unlicensed spectrum should also be sought, which for enhanced management purposes should be done by an integrated mechanism. The geolocation database approach of TVWS provides a good basis on which also to do this, providing the necessary tools to achieve resource usage management of unlicensed devices themselves, not only in TVWS, but also in conventional unlicensed bands. This coordination could be achieved by either the same database(s) that are operating in a trusted form under the certification of the regulator (as in TVWS), or another database that does not need to be trusted as there are no legal implications if mistakes are made in unlicensed spectrum (as long as the decisions and effects stay within the scope of the unlicensed spectrum and associated band rules). It is important to note, however, that such an untrusted geolocation database cannot be responsible for management of opportunistic secondary spectrum access as would have to be implemented by the regulator or qualified by the regulator, or for sharing among licensed users.

Using TVWS databases and the Ofcom TVWS Pilot in the UK as one pertinent example, a key observation is that the geolocation databases and the entire underlying framework are seen as a baseline that can used and extended to a range of forms of opportunistic spectrum access, also in other bands, and more generally for the imposition by the regulator of locally-optimised regulations based on geolocation and other (e.g., technical) information. This is essentially what the geolocation databases already do in a simple form in the TVWS case, varying the allowed power on a per-channel basis for each given location. The key is that to allow opportunistic secondary spectrum access, the regulator has to be involved and the geolocation databases and framework are a good automated model one to take forward the concept with the involvement of the regulator. Moreover, as has been borne out in the LSA case for example, it can even be necessary (or even beneficial) for the regulator to be involved in cases where there is a direct agreement between the incumbent spectrum user and the opportunistic user. Given all of the above, geolocation database functionality must either exist within the regulator, or be approved by the regulator and implemented by a trusted party. However, in the UK TVWS case, for example, some geolocation database calculations are done solely at the regulator for information privacy reasons, and some more computationally challenging ones (e.g., protection of PMSE services) are done by the trusted database outside of the regulator’s scope. Hence, in this particular case, the workload is split between the regulatory and the (trusted) non-regulatory domains, and geolocation databases exist in both domains.

Given such observations on the capabilities of geolocation databases, it is suggested here that such an integrated system for spectrum coordination supporting 5G is based on them, due to: (i) the need to involve the regulator and coordinate heterogeneous spectrum usage/sharing in a way that the regulator can trust, (ii) the establishment and trialling of them in various contexts and likely further building on such concepts in regulatory and other circles, and (iii) the advancement of such approaches (e.g., LSA) to support mobile communications cases. Moreover, it is suggested that these databases should be of three forms: (i) a form of database in the presence of and run by the regulator, (ii) a form of database trusted and likely certified by the regulator but existing and being run outside of the scope of the regulator, and (iii) a final (optional) form of database that is untrusted, and existing and being run outside of the regulator.
III.  EXAMPLE HIGH-LEVEL ARCHITECTURAL VIEW

In view of the observations in previous sections, an initial potential viewpoint on the architecture for such a management system is given in Figure 1. The geolocation databases within this architecture are depicted with a green background. They first comprise a Regulatory Geolocation Database (RGD), operating within the scope of the regulator. This is because it is essential that some information originates and stays only within that scope, and there may be privacy or other concerns that limit certain calculations to only being done inside of that scope, as described in Section II. Outside of the regulator’s scope there are the Trusted Non-regulatory Geolocation Database (T-NGD) and Untrusted Non-regulatory Geolocation Database (U-NGD). These entities take the complexity of more advanced resource sharing calculations out of the regulator’s direct implementation. The T-NGD serves cases where a very high guarantee of the result of the calculations or other trustworthiness is necessary, for reasons as described in Section II. The U-NGD may be able to handle cases where no guarantee of trustworthiness or reliability is necessary, e.g., in unlicensed spectrum should a number of unlicensed terminals choose to operate under the control of such a database but where if the database makes an error, for example, no liability implications will result. It is useful to have a separate such entity, as there is a large overhead in managing a T-NGD, including achieving certification by the regulator, and also associated implied workload at the regulator. Hence, there will be a practical limit on the number of T-NGDs that can be deployed.

Certain aspects of this proposal placed within a generalized 5G network are also depicted, strongly inspired by an author’s prior work on and experience of the IEEE 1900.4 heterogeneous network/resource management standard [8]. This is with a pink background in Figure 1. Here, there are entities on both the network side (NRM) and terminal side (TRM) that deal with management through policies (the network side having overall control, but leaving some aspects for terminal TRM local decisions through the policies the NRM creates). On both sides, there are also entities that implement/enforce decisions (NRC and TRC), and obtain context information to feed back to the NRM and TRM (NRIC and TRIC) to assist in making the correct decisions based on the situation. There are of course separate RANs, for which decisions must be made and through which signalling between...
the network side and terminal side may take place. Further, there may be a separate signalling carrier operating outside of the mobile network, also depicted in Figure 1.

The terminals/devices operating within the unlicensed domain are illustrated with a blue background in Figure 1. Importantly, they will generally interact with the U-NGD, which could manage aspects such as better coordination of their unlicensed spectrum usages, taking into account the resource requirements of the devices and also the capabilities of the devices to aggregate resources. Further, the T-NGD might also manage this domain, e.g., should it wish to pair licensed and unlicensed resource usages such as is currently done in LTE-U/LAA or example.

It is noted that as well as managing allowed spectrum access in the sense of avoiding interference with incumbents, spectrum access among the secondary/opportunistic spectrum users might also be managed. In an advanced implementation scenario, this can take into account the resource requirements of the secondary/opportunistic users and allocate the spectrum accordingly, also taking into account the bandwidths that devices will need and the need to aggregate resources (e.g., channels) to achieve necessary capacity. Further, under an even more advanced scenario, such a database could play the role of pairing devices and their traffic requirements with link opportunities that are available in local areas, including the optimal allocation of which opportunities should be aggregated to achieve requirements.

**IV. PERFORMANCE ASSESSMENT, AND OBSERVATIONS ON THE APPLICATION TO 5G IN TV WHITE SPACE**

The first author of this paper is leading a major trial of TVWS technology within the UK, intended to verify the operation and performance of the UK/EU TVWS framework and the associated white space devices (see, e.g., [9], [4]). The UK/EU framework reflects the implementation of the RGD and T-NGD in Figure 1, noting that the framework could (if it was required) also handle the management of resource allocations among the secondary/opportunistic users, without changes, aside from enhancements to the RGD and T-NGD internal processing. Further, with advancements to the information sets supported and the development of allocation policies, it could also implement the obtaining of additional context information (e.g., network/terminal traffic requirements) and the allocation of resources as well as aggregation of resources to match those requirements.

Given the parallels described above, the architecture we propose in this paper maps directly to the UK/EU TVWS framework, particularly the RGD and T-NGD as introduced in Section IV, the purpose being to authorise opportunistic access of another services’ spectrum by 5G networks and devices. Conceptually however, the work and assessment we present here can also be seen as the framework controlling a 5G operator opportunistically using the incumbent 5G operator’s currently unused spectrum, via an agreement with the incumbent operator and the associated trusted database (T-NGD) operating within the scope of the incumbent operator. In this context, we concentrate on two scenarios which quite closely map to mobile communication scenarios: (i) a macro-cell case where the chosen propagation model particularly applies to the downlink, and (ii) a 5G small-cell case. The

<table>
<thead>
<tr>
<th>Table 1: Performance Assessment Configuration Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Macro-cell (downlink)</strong></td>
</tr>
<tr>
<td>Trans. Ht. (m)</td>
</tr>
<tr>
<td>30</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

Fig. 2. The assessed London “M25” area.

![Number of available contiguous channels for the macro-cell (downlink) scenario vs. location within the London “M25” area.](image)

Fig. 3. Number of available contiguous channels for the macro-cell (downlink) scenario vs. location within the London “M25” area. Class 3 device.

Moreover, we assess the London, UK “M25” area—which encompasses the entire London area stretching quite significantly out into the suburbs and surrounding countryside. Results are obtained via our implementation of a device that complies with the UK/EU framework and queries a geolocation database operating in the UK/EU framework (see
at multiple locations to obtain and further compute a range of statistics on TV channel availability and resulting capacity under our scenarios. For all results, the entire London “M25” area as depicted in Figure 2 is queried, in equal steps of 0.01 degrees in latitude and longitude. This gives a total of 2,775 locations/values for each of the sampled data sets.

First, in Figure 3, we choose one representative case that geographically maps contiguous channel availability for a Class 3 device under the macro-cell scenario. It is noted that the variation in availability here is reasonably representative of what can be seen for geographical mappings of statistics presented in other figures in this paper, although it is noted that the small cell scenario generally shows far higher availability/capacity and a far lower geographical variation. Noticeable here is that there are distinct areas in which there is consistently good availability, and others where availability is reduced. This is caused by TV transmission patterns in the London area [9]. Moreover, in the areas of good availability, there are still “bloches” of significantly reduced availability. These are caused by localised PMSE services. A key observation given all of this is that the 5G geolocation database (i.e., T-NGD) could manage the resources it attributes with a reasonable degree of certainty and stability for each given location, even through the performances among the different locations vary significantly.

Considering the number of available channels for the five classes of spectrum mask under the UK/EU framework [4], depicted in Figures 4 and 5, it is very clear that for the macro-cell case the probability of a least a given number of channels being available decreases reasonably linearly as that number of channels is increased. For the small cell case the number channels is consistently quite high. However, if those channels have to be contiguous, then the reduction is a lot more severe. Moreover, under the contiguous channels case there is very little difference between the performances of the various spectrum mask classes in terms of channel availability, and in the small cell case there is virtually no difference. This implies that a contiguous channel device (e.g., using three or more channels) for 5G in TVWS could be architected to have a relatively poor spectrum mask with very little loss in allowed transmission EIRP. For the small cell case, such a device would be reasonably viable in terms of being able to aggregate a large number of channels (see Figure 5(b)), where the good performance of relatively poor spectrum masks chimes well with the need to reduce production costs for small cells and address technical challenges. Further, from both these figures it is clear that there is a very significant advantage in striving for devices being able to use multiple non-contiguous channels, such as based on a Filter-Bank Multi-Carrier waveform, for example. Finally, referring to Figure 6, it is clear that the situation in terms of number of available channels is reflected in the capacity that is achievable by aggregating all channels at maximum allowed EIRP on a per-channel basis. However, in this case it is important to note that EIRPs are often higher than the 30 dBm and 20 dBm thresholds assumed in other figures; otherwise there would be a direct proportional mapping of number of available channels to the achieved rate.

A clear observation, as would be expected in an unlicensed spectrum access scenario and particularly in TVWS where the
allowed EIRP of the unlicensed access varies on a per-location and per-channel basis, is that there is vast variation in what can be achievable in given locations and channels. This emphasizes the need for the architecture to be robust against achievable capacity and likely also reliability variations. Using such observations and concentrating on the use of the wireless link for 5G end-user access (e.g., the abovementioned macro-cell and small cell cases) a proposed signalling chart for TVWS access to cover such cases, compliantly with the UK/EU TVWS framework, is as in Figure 7. This would be implemented between the NRM and T-NGD in Figure 1, i.e., over the T-N interface.

V. CONCLUSION

In this paper, we have argued for the introduction of a geolocation database-based architecture to enhance heterogeneous spectrum usage and spectrum sharing for 5G. The key purpose of this is to realise capacity (hence spectrum) requirements for 5G services, particularly at lower frequencies. The architecture is intended to enable the opportunistic usage of another service’s spectrum by 5G systems, sharing among 5G networks/operators and devices, better coordination of unlicensed spectrum sharing in 5G, and optimal dynamic allocation of resources dependent on traffic loads in a range of scenarios. Finally, this paper has assessed the concept through an analogy and study of results obtained in a major trial within the Ofcom TV White Spaces Pilot, led by the authors. It has derived further observations for the 5G architecture based on these results, including particularly the need to take into account and balance resources based on the uncertainty in spectrum access that persists in TVWS cases.

Fig. 6. CCDFs of achievable capacity for the London “M25” area aggregating all available channels at maximum allowed EIRP on a per-channel basis: (a) macro-cell (downlink) scenario, (b) indoor small cell scenario.

Fig. 7. Signalling diagram for 5G macro-/small cell access to TVWS, derived based on TVWS performance observations in this paper.

ACKNOWLEDGEMENTS

This work has been supported by the H2020-ICT-2014-2 project 5G NORMA. The authors would like to acknowledge the contributions of their colleagues in NORMA, although the views expressed are those of the authors and do not necessarily represent the project. This information reflects the consortium’s view, but the consortium is not liable for any use that may be made of any of the information contained herein. This work has also been supported by the “Spectrum Overlay through Aggregation of Heterogeneous Dispersed Bands” project, ICT-SOLDER, www.ict-solder.eu. The authors would like to particularly thank Heikki Kokkinen of Fairspectrum.

REFERENCES


