The RF Spectrum-Sharing Challenge

Participant Guide

RFSSC Organizers

May 15, 2024

Version 4

Contents

1 Introduction

This document describes an *RF Spectrum Challenge* aimed at spurring innovative technologies for automatic RF signal detection, characterization, and recognition. By *detection* we mean presence detection; by *characterization* we mean estimation of key parameters such as symbol rate, pulse-repetition rate, and carrier frequency; and by *recognition* we mean system or signal classification. The problem is meant to encompass a much wider set of possible RF scenarios than do the usual academic modulation-recognition problem settings. This will be accomplished by including a large number of signal types, a wide range of signal parameters, a wide range of signal-to-noise and signal-to-interference-and-noise ratios (SNRs and SINRs), cochannel interference situations, intermittent signals, and signals not represented in any dataset supplied to the participants by the organizers.

To entice world-class researchers and algorithm developers to participate in the Challenge, we will wrap the basic technical goals in a future-looking *spectrum-sharing* problem in which a limited RF spectral band is to be shared by many different kinds of users over time and space. The nature of the sharing problem will force the Challenge participants to achieve the stated technical goals without unduly constraining the nature of the potential solutions.

2 Spectrum-Sharing Background

In existing and proposed spectrum-sharing systems, conflicts are partially mediated by creating distinct classes of users that have different access rights to the shared band of frequencies. For example, Federated Wireless shows the CBRS sharing scheme illustrated in Figure 1. The priority (or privileged access level) for the three user classes decreases going downward from the tip of the user triangle. The military users have highest priority, the licensed users (payers) have second priority (yielding to the military users as needed), and the general users have the third and lowest priority, having to yield to both the military and licensed users as needed.

The ability to rapidly and accurately sense the spectrum to determine if users with higher priority are present is a key to the success of this kind of spectrum-sharing system. From the Federated Wireless white paper,

> ***A critical element of spectrum sharing architecture is the Environmental Sensing Capability (ESC). When an ESC sensor detects a federal transmission, it activates a protection zone and informs the SAS to dynamically reallocate users in the area to other parts of the band. The ESC is a key component of the system since it opens***

---

[1] http://federatedwireless.com/wp-content/uploads/2017/02/CBRS-Spectrum-Sharing-Overview.pdf.

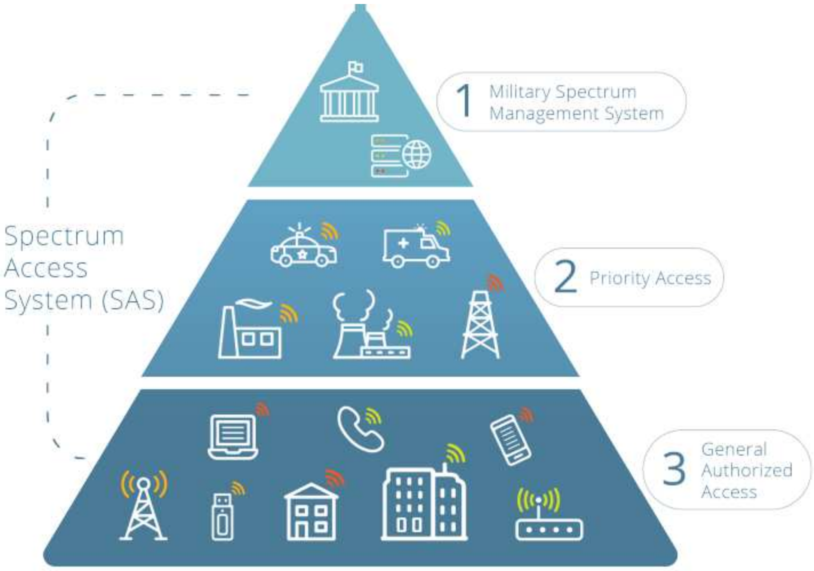

Figure 1: Illustration of real-world proposed Citizens Broadband Radio Service shared spectrum user tiers (from Federated Wireless' spectrum-sharing whitepaper[1]).

> ***the value of shared spectrum along U.S. coastal counties where over 50 percent of the population lives according to NOAA. By using many sensors in a network, the ESC can assure high availability of wireless spectrum.***

In the RFSSC, we adopt a somewhat more flexible and challenging spectrum-sensing system. There are still three tiers of users, but the rules governing their behavior are more complex. This complexity permits us to create a wide variety of sharing-compliant scenarios that will require new sensing methods to correctly assess. In the RFSSC system, there are *primary*, *secondary*, and *urgent* users. In some sense, the urgent users correspond to the CBRS military users, the primary users correspond to the CBRS licensed users, and the secondary users correspond to the lowest-priority CBRS general users.

The RFSSC sharing rules, user classes, and corresponding signal types are described in the following sections.

3 Spectrum-Sharing Problem Description

The spectrum-sharing problem in the Challenge must force the participants to detect, characterize, and recognize (DCR) all signals in an organizer-provided test setting. The test setting is anticipated to consist of a sequence of data files rather than an over-the-air setting involving transmitted signals. This will greatly reduce the organizers' difficulties in running the Challenge.

The problem is stated in terms of a hierarchy of sharing privileges and rules involving three kinds of users: primary, secondary, and urgent. The primary users have right-of-way relative to the secondary users. The urgent users have right-of-way relative to both primary and secondary users, but can access the shared RF band only for strictly limited amounts of time and numbers of accesses.

The RF band that is shared is divided into N_c adjacent channels of width B_c, so that the total bandwidth of the sharing band is $N_c B_c$ Hz. An urgent user is allowed N_a accesses of the channels in T_u seconds. The overall amount of time that the urgent user is allowed to use the channels in T_u seconds is limited to $T_a < T_u$ seconds. Once the urgent user has exceeded N_a uses in T_u seconds or has occupied the channel(s) for more than T_a seconds in T_u seconds, the user is no longer urgent and can be interfered with and primary and secondary users can occupy channels adjacent to the formerly urgent (nonurgent) user. All of these urgent-user complications force Challenge participants to perform DCR on all signals over all times, and to keep a count of the appearances of each urgent user. An urgent user that is no longer urgent is called a nonurgent user and must follow the rules for secondary users for the remaining time in the current T_u-second interval.

The primary and secondary users are defined for the participants. The definition will include the system name (if appropriate), the modulation type, and the ranges of all relevant modulation parameters. This will allow participants to develop their own training (for ML) and testing (for ML and DSP) datasets.

To describe the sharing setup, three key notions require definition.

- **Definition 1: N-Adjacency.** Two signals are N adjacent if there are N or fewer frequency channels of width B_c separating the edges of their occupied bands. For signals with strictly bandlimited spectra, such as square-root raised-cosine PSK/QAM, N-adjacency can be checked by mathematical analysis since the occupied bandwidth is easily computable. For other signals, we define the occupied bandwidth to be the 99% bandwidth.
- **Definition 2. An Occupied Channel.** A frequency channel of width B_c Hz that contains total noise power of P_n and total non-noise power of P_s is unoccupied if $P_n/P_s > 100$, else it is occupied.
- **Definition 3. A Channel Access.** A *channel access* is defined as a transmission during a subepoch on a spectral channel.

The basic spectrum-sharing problem is described next.

1. The RF sharing band is segmented into N_c adjacent, non-overlapping channels (or *spectral channels*) of B_c Hz. The channel parameter B_c is expected to be much smaller than the smallest expected occupied bandwidth over all primary, secondary, and urgent users' signals.

2. The sharing experiment lasts for T seconds which is segmented into T/T_u adjacent, non-overlapping sharing-time epochs of T_u seconds. Each of these epochs is further finely divided into N_d subepochs. Scoring (described in Section 5) will make use of the subepochs.
3. The number of distinct primary-user signal types is N_p and the number of distinct secondary-user signal types is N_s. These numbers include all possible allowed modulation-parameter variations for the two user classes.
4. Users can transmit signals that occupy an integer number of channels if allowed to access those channels using the following rules.
5. During the ith epoch:
   (a) Primary users can occupy any channel that is occupied by a secondary user.
   (b) Secondary users must evacuate a channel if a primary or urgent user begins to transmit on that channel.
   (c) Primary users can occupy any unoccupied channel that is not M_p-adjacent to an urgent user.
   (d) Secondary users can occupy any unoccupied channel that is not M_s-adjacent to an urgent user.
   (e) Urgent users can occupy any channel, with the expectation that other users will quickly evacuate, provided their in-epoch transmit time has not exceeded T_a seconds and that the number of in-epoch channel accesses is less than N_a. Otherwise, the urgent user might occupy any channel but cannot expect the channel will remain interference-free.
   (f) Both primary and secondary users can occupy a channel that is occupied by a nonurgent user. A nonurgent user is a formerly urgent user that has exceeded the in-epoch transit time or number of in-epoch channel accesses.
   (g) Both primary and secondary users can occupy a channel that had been previously occupied by an urgent-user radar signal only if that radar signal has been absent for a period of time exceeding twice the signal's PRI and the urgent-user radar signal had not yet become nonurgent.
   (h) A secondary user can occupy a channel that had been previously occupied by a primary-user radar signal only if that radar signal has been absent for a period of time exceeding twice the signal's PRI.

These sharing rules and user-type definitions allow the organizers to create both simple and highly complex RF environments, which are used to spur the development of innovative effective

DCR technologies. By allowing the urgent-user class to remain unspecified (or vaguely specified), the organizers evade the typical machine-learning approach involving massive datasets containing all the expected signals–the urgent-user class cannot be fully anticipated. Therefore machine-learning participants will have to directly address the open-set (or dataset-shift or generalization) problem and the signal-processing participants will have to use blind or partially blind detectors and estimators.

4 Involved Signal Types and Parameters

The RFSSC will employ a universe of signal types and associated parameters that provide a tradeoff between realism and practicality of generation. Basic realism can be obtained by using modulation types that are bandwidth efficient and contain pseudo-MAC-layer components, such as framing and repeated frame preambles. These are *pseudo* because we are not developing modulators **and** demodulators for these signals; we are not developing or using any actual communication systems. But the presence of realistic elements such as framing, pilot tones, periodically repeated synchronization sequences, etc., can significantly change the statistical nature (and appearance of ML mainstays such as spectrograms) of an otherwise textbook signal.

4.1 Communication Signals

The following signal types and parameter ranges are used for primary-user communications transmitters.

1. OFDM.
   - (a) Captured LTE, 5 MHz occupied bandwidth (OBW).
   - (b) Generic, 512 subcarriers, 14 identical pilots, QPSK modulation, CPs of 1/4, 1/8, 1/16, and 1/32, OBW 10 MHz.
   - (c) Pilots are on subcarriers [zero indexing] 42, 74, 106, 138, 170, 202, 234, 266, 298, 330, 362, 394, 426, and 458.
   - (d) The pilot sequences are randomly chosen per signal instance, are identical in each pilot subcarrier, and have length 32.
   - (e) The subcarrier power levels are chosen randomly per signal instance, and can vary over a 3-dB range.
2. PSK/QAM. For all PU communication signals that employ repeated known synch sequences, the sequence is chosen randomly per signal instance and is used throughout that instance.

   (a) BPSK, QPSK, and 8PSK. Symbol rates in the set $\{1, 2, 3, 4, 5\}$ MHz. Square-root raised-cosine pulses with rolloff of 0.3. Every P symbols transmit a known Q-symbol sequence to aid the hypothetical receiver in synchronization, channel estimation, and user identification. $P = 1000$ and $Q = 50$.
   (b) 16QAM and 64QAM. Standard constellations. Symbol rates in the set $\{1.5, 2.5, 3.5, 4.5, 5.5\}$ MHz. Square-root raised cosine pulses with rolloffs of $\{0.2, 0.35, 1.0\}$. Every P symbols transmit a known Q-symbol sequence to aid the hypothetical receiver in synchronization, channel estimation, and user identification. $P = 1000$ and $Q = 50$.
   (c) GMSK. Symbol rate of 5.0 kHz. No periodically transmitted synch/pilot/ID sequence.

The following signal types and parameter ranges are used for secondary-user communications transmitters.

1. DSSS. For all SU communication signals that employ repeated known synch sequences, the sequence is chosen randomly per signal instance and is used throughout that instance.
   (a) DSSS BPSK and DSSS QPSK. Chip rates in the set $\{1, 2, 3, 4, 5\}$ MHz. Processing (spreading) gains in the set $\{31, 63, 127, 255\}$. Square-root raised-cosine pulses with rolloff of 0.3 and MLSR sequences. Every P data symbols transmit a known Q-data-symbol sequence to aid the hypothetical receiver in synchronization, channel estimation, and user identification. $P = 2000$ and $Q = 75$.
2. CPM.
   (a) MSK and GMSK. Symbol rates anywhere in the interval $[100, 500]$ kHz. For GMSK $BT = 0.3$. Every P symbols transmit a known Q-symbol sequence to aid the hypothetical receiver in synchronization, channel estimation, and user identification. $P = 1500$ and $Q = 75$.
3. Narrowband FSK.
   (a) 2FSK and 4FSK. Symbol rates in the set $\{2.4, 3.6, 4.8\}$ kHz. No periodically transmitted synch/pilot/ID sequence. Carrier-phase-coherent FSK.
4. PSK/QAM.
   (a) BPSK, QPSK, and 8PSK. Symbol rates in the set $\{1.1, 2.1, 3.1, 4.1, 5.1\}$ MHz. Square-root raised-cosine pulses with rolloffs of 0.25. Every P symbols transmit a known Q-symbol sequence to aid the hypothetical receiver in synchronization, channel estimation, and user identification. $P = 2000$ and $Q = 50$.

5. FH.

    (a) 40 hop frequencies, hop-frequency separation of 1.25 MHz, hop rate of 1.9 kHz, hop duration of 0.5 ms, modulation on-hop of BPSK, symbol rate of 1.0 MHz, bits per hop of 500. No periodically transmitted synch/pilot/ID sequence.

The signal types and parameter ranges for urgent-user communication transmitters are withheld from Participants.

4.2 Radar Signals

The following signal types and parameter ranges are used for primary-user radar transmitters.

1. PU_RADAR_1: Pulsed, unmodulated, two pulse repetition intervals (PRIs), two pulse widths, two carrier offsets.
2. PU_RADAR_2: Pulsed, unmodulated, PRI of 2.577 ms, pulse width of 4.93 μs.
3. PU_RADAR_3: Linear FM (LFM), PRI of 865 μs, pulse width of 11 μs.

The following signal types and parameter ranges are used for secondary-user radar transmitters.

1. SU_RADAR_1: Pulsed, maximal-length sequence (MLS) modulation with 255 symbols per pulse, PRI of 1.103 ms, and a pulse width of 94 μs.
2. SU_RADAR_2: Pulsed, unmodulated, variable frequency offset, variable PRI in the interval $[3.019, 3.071]$ ms.
3. SU_RADAR_3: Pulsed, Barker-13 pulses, PRI of 1.1 ms, and a pulse width of 120 μs.
4. SU_RADAR_4: Pulse-Doppler, unmodulated, multiple PRIs, pulse width of 1.04 μs.

The signal types and parameter ranges for urgent-user radar transmitters are withheld from Participants.

4.3 Rationale for Signal Choices

The overarching rationale is to permit the organizers to develop a sequence of Challenge Levels that are increasingly difficult to solve using conventional or even recent cutting-edge signal-processing or machine-learning techniques–we want to spur innovation. This means that we must be able to introduce signals and signal combinations that are ambiguous relative to those conventional or cutting-edge techniques. We must have superficial similarity between some pairs of primary- and urgent-user signals, and between some pairs of secondary- and urgent-user signals. We must be able to construct cochannel signals that are in principle jointly detectable.

This rationale leads to the specification of similar but disjoint signal sets for the three classes of users. For easy Challenge Levels, we can pick representatives from the three sets that are maximally distinct. For moderate Challenge Levels, we pick mostly distinct signals, but a few that are quite similar. For difficult Challenge Levels, we pick mostly similar signals from the sets so that distinguishing among the different signals is maximally difficult, requiring new solutions relative to state-of-the-art.

5 Metrics for Detection and Characterization

5.1 Definition

Typical metrics for this kind of Challenge are probabilities of detection and false alarm for presence detection and mean-squared error for parameter estimates such as symbol rate and PRI. However, this kind of scoring is complex and challenging for the organizers. Instead, we propose a kind of metric that is appropriate for a spectrum-sharing problem: ***ability to transmit.*** For each subepoch in each epoch, and for each channel, the participant will have to provide a two-element binary vector $[P\ S]$. The first element denotes the ability of a primary user to transmit on the given channel at the subepoch start time. The second element denotes the ability of the secondary user to transmit on the given channel at the subepoch start time.

These *ability-to-transmit* values can only be correct over time if the RF scenario is continuously correctly assessed. Due to the imposed sharing rules, this assessment must take into account the arrangement in time and frequency of all secondary, primary, urgent and nonurgent (formerly urgent) users in each epoch. The ability of a primary to transmit on any given channel is contingent on the number and location of any urgent users in the epoch. To determine whether a signal belongs to an urgent user or a nonurgent user, the behavior (channel accesses and total transmission time) of the urgent user must be closely tracked. For the secondary users, the ability to transmit requires the urgent-user knowledge that the primary users must cultivate, but also similar information about the primary users. In this way, the simple binary ability-to-transmit pairs $[P\ S]$ entail a large number of inferences regarding the spectral situation, but avoid complex scoring schemes involving comparing participants' sequences of parameter estimates and decisions over time and frequency.

The organizers will have created the correct set $\{[P_{i,j,k}\ S_{i,j,k}]\}$, where i denotes the epoch index, j denotes the subepoch index, and k denotes the channel index. This simple set will be compared elementwise to each participant's submitted set. The participant's score is the number of matching elements, or a normalized version of that count.

5.2 Examples

Primary Users, Secondary Users, but No Urgent Users.

First consider the case where no urgent users are used, and this is known by the participants. By the rules in Section 3, all $P_{i,j} = 1$. The values of $S_{i,j}$ will depend on whether there is a primary user present in the channel, or in the channels on either side. Therefore, correct answers for $S_{i,j}$ will depend on the ability of the participant to correctly and reliably detect all primary users in the test data. Since primary and secondary users do not need to check for the presence of secondary users before transmitting, this situation does not require that the participant can correctly and reliably detect any of the secondary users.

Primary Users, No Secondary Users, A Single Urgent User.

Next consider the case where no secondary users are used, a single urgent user can appear in the test, and this is known to the participants. Now the ability to produce correct P and S values depends on the ability to distinguish between secondary and urgent users, as well as to count the number of channel accesses by the urgent user and to keep track of its total transmission time within each epoch. It is not required to be able to perform DCR on each and every secondary user type, but only that the urgent and primary users' signals can be reliably recognized. A failure to properly count the urgent-user's channel accesses or transmit time will translate into erroneous reported P and S values. Failure to properly detect and recognize all primary users will translate into erroneous S values.

Primary Users, Secondary Users, and A Single Urgent User.

Now consider a scenario in which we have primary and secondary users, and a single urgent-user signal type that can appear in the test, and this overall situation is known to the participants. The ability to produce correct P and S values depends on the ability to correctly detect and distinguish all three kinds of users. For the primary user (getting P correct), the secondary and urgent-users' signals must be distinguished, and the urgent-user's signal must be correctly tracked over time to determine when it becomes nonurgent. For the secondary users (getting S correct), the primary and urgent-users' signals must be distinguished, and the urgent-user's signal must be tracked over time to find the time instants at which is becomes nonurgent.

Primary Users, Secondary Users, and Multiple Urgent Users.

In the most complex scenario, there are multiple primary, secondary, urgent, and nonurgent users. Recall that a nonurgent user's signal is distinct from all primary and secondary signal types, and has persistent sufficiently long or appeared sufficiently many times in some time window to render it nonurgent. The participants are provided the information that all primary and secondary users' signals can come and go, and that any number of urgent-user signals can appear. Urgent-user signals can appear cochannel with an existing primary, secondary, urgent, or nonurgent user.

For the primary user (getting P correct), the secondary, urgent, and nonurgent signals must be distinguished, and this may involve correct identification of each of a number of cochannel signals. For the secondary user (getting S correct), the primary, urgent, and nonurgent signals must be distinguished, and again cochannel situations might be common.

These challenges are made more difficult when urgent and nonurgent signals are spectrally identical to primary or secondary signals. (See Section 4.)

5.3 Scoring

Several scores have been developed to assess the overall RF scene-analysis performance of the Participants. The scene analyses they perform lead to modulation-type decisions and associated numerical signal-parameter estimates, such as estimates of symbol rate, chip rate, carrier frequency, SQRC filter rolloff, and SNR. The Participants will use these estimates, as obtained over an entire Organizer-provided scene data file, together with the Organizer-provided sharing rules to create the ability-to-transmit values for Primary and Secondary users.

5.3.1 Alternative Scores

The time-frequency plane is divided into frequency channels, time epochs, and time subepochs. There is an ability-to-transmit value (zero or one) for each pair of frequency channel and time subepoch and for each of the Primary and Secondary user classes–this implies a matrix with rows corresponding to subepoch times and columns corresponding to frequency channels.

Let's call the true ability-to-transmit matrix for Primary users P, which has N rows and M columns; denote the product by $E = NM$. The true ability-to-transmit matrix for Secondary users is S, which has the same dimensions as P. A participant provides two matrices $\hat{P}$ and $\hat{S}$, also having the same dimensions as P. The scoring for Participants' submissions is based on one or more mathematical comparisons between the true matrices and the submitted ones. Here are some scoring definitions:

1. **Basic Sharing.** This score lies on $[0, 1]$, with a score of 1 representing perfect sensing.

$$
\begin{aligned}
P_d &= P - \hat{P} && (1) \\
S_d &= S - \hat{S} && (2) \\
T_0 &= \frac{1}{2}\left[\left(1 - \frac{1}{E}\sum_{i,j} |P_d(i,j)|\right) + \left(1 - \frac{1}{E}\sum_{i,j} |S_d(i,j)|\right)\right] && (3)
\end{aligned}
$$

The Primary and Secondary Basic Sharing scores are simply the values enclosed by the parentheses in (3).

2. **Conditional Sharing Type 1.** Let the set A denote all those (i, j) values for which $P(i, j) = 0$. That is, those locations in the time-frequency plane for which the Primary user is not allowed to transmit. Let the set B denote all those elements in A for which $P_d(i, j) < 0$. That is, those locations where the Primary is not allowed to transmit, but the Participant declares the Primary *can* transmit. Then the Type-1 Conditional Sharing score for Primary users is
$$T_1(P) = 1 - \frac{|B|}{|A|} \tag{4}$$
Similarly, define analogs to the matrices A and B for the Secondary user class and call these C and D. Then the Type-1 Conditional Sharing score for Secondary users is
$$T_1(S) = 1 - \frac{|D|}{|C|} \tag{5}$$
and the overall Type-1 Conditional Sharing score is
$$T_1 = \frac{1}{2} \left(T_1(P) + T_1(S) \right). \tag{6}$$

3. **Conditional Sharing Type 2.** Here we condition on those locations in the time-frequency plane where the truth ability-to-transmit values indicate transmission is possible, but the Participant decided it is not. Let A denote all those (i, j) values for which $P(i, j) = 1$ and B denote all those elements in A for which $P_d(i, j) > 0$). Form similar sets C and D for the Secondary users and S and S_d. Then the Type-2 Conditional Sharing scores are
$$\begin{aligned} T_2(P) &= 1 - \frac{|B|}{|A|} &(7)\\ T_2(S) &= 1 - \frac{|D|}{|C|} &(8)\\ T_2 &= \frac{1}{2} \left(T_2(P) + T_2(S) \right). &(9) \end{aligned}$$

4. **Basic Occupancy.** The occupancy-based scores assume that the Participant provides occupancy matrices for Primary, Secondary, and Urgent users, denoted by $\hat{P}_O$, $\hat{S}_O$, and $\hat{U}_O$. The corresponding truth matrices are P_O, S_O and U_O. These matrices have elements that are either zero (time-frequency cell is unoccupied) or one (it is occupied). As with the ability-

to-share matrices, we form the difference matrices

$$\begin{aligned} P_{O,d} &= P_O - \hat{P}_O & (10) \\ S_{O,d} &= S_O - \hat{S}_O & (11) \\ U_{O,d} &= U_O - \hat{U}_O & (12) \end{aligned}$$

The Basic Occupancy score is then the average number of elements where the difference between truth and submission is non-zero,

$$T_3 = \frac{1}{3}\left[\left(1 - \frac{1}{E}\sum_{i,j}|P_{O,d}|\right) + \left(1 - \frac{1}{E}\sum_{i,j}|S_{O,d}|\right) + \left(1 - \frac{1}{E}\sum_{i,j}|U_{O,d}|\right)\right] \tag{13}$$

5. **Conditional Occupancy.** This score measures the number of occupancy errors for a user class conditioned on the total number of true occupancy cells for that class. Let Y denote the number of elements of P that are equal to one (the total occupancy, measured in time-frequency cells, for the Primary users) and Z denote the set of (i, j) values such that $P_{O,d}(i, j)$ differs from zero (number of erroneous Primary-user occupancy decisions made by the Participant). Then the Conditional Occupancy score for the Primary user class is

   $$T_4(P) = 1 - \frac{|Y|}{|Z|}, \tag{14}$$

   provided Z is not empty. The Conditional Occupancy scores for the Secondary and Urgent user classes, $T_4(S)$ and $T_4(U)$, are defined analogously.

6. **Overestimated Occupancy.** This occupancy-based score looks at how much the Participant overestimated the bounding boxes of each user class. Again we define the score for the Primary user class, but the definition extends easily to the other two classes. Define Y as before. Define W as the set of (i, j) values such that $P_{O,d}(i, j)$ is less than zero. These elements indicate that the truth value is zero (Primary user is not present at that time-frequency cell) but the Participant's Primary-user occupancy matrix is positive (Participant thinks a Primary user is indeed present). Then the Overestimated-Bounding-Box Occupancy score for the Primary users is

   $$T_5(P) = 1 - \frac{|Y|}{|W|}, \tag{15}$$

   provided W is not empty. The Overestimated Occupancy for the Secondary and Urgent user classes, $T_5(S)$ and $T_5(U)$, are defined analogously.

7. **Underestimated Occupancy.** Here, we devise a score that assesses how badly bounding

boxes are underestimated. Define Y as before. Let X equal the set of (i, j) values such that $P_{O,d}(i, j)$ is greater than zero. These elements indicate that the truth value is one (Primary user is present at that time-frequency cell) but the Participant's Primary-user occupancy matrix is zero (Participant thinks that the Primary user is not present). Then the Underestimated-Bounding-Box Occupancy score for the Primary users is

$$T_6(P) = 1 - \frac{|Y|}{|X|}, \tag{16}$$

provided X is not empty. $T_6(S)$ and $T_6(U)$ are defined similarly.

8. **Simple Occupancy.** The previous alternative scores assumed that the Participant made an attempt to characterize each detected signal (each distinct bounding box in the time-frequency plane) as belonging to one of the three user classes. Therefore, there are three different occupancy maps. It may be the case that a Participant fails to make these classification decisions, or makes them poorly. But that same Participant may do well on determining all the true bounding boxes in the plane. The final alternative score is based on a simple occupancy map that does not distinguish between the user classes. The true simple occupancy is the binary matrix O_O, which has the same dimensions as P_O, S_O, and U_O. A Participant that does make user-class decisions can produce $\hat{O}_O$ as the union of $\hat{P}_O$, $\hat{S}_O$, and $\hat{U}_O$.

   The difference between the true simple occupancy and the estimated simple occupancy is

$$O_{O,d} = O_O - \hat{O}_O. \tag{17}$$

   The score for simple occupancy is then

$$T_7 = 1 - \frac{1}{E} \sum_{i,j} |O_{O,d}| \tag{18}$$

   If the submitted simple occupancy matrix matches the true simple occupancy matrix at every matrix position, then $T_7 = 1$. If the submitted simple occupancy matrix is different from the true simple occupancy matrix at every matrix position, then $T_7 = 0$.

5.3.2 Final Score

The alternative scores in Section 5.3.1 will be used to create a single final score (a single number) to characterize a Participant's submission. The final score should be sensitive to small differences in two Participants submissions so that they may be distinguished even though both may be very close to the ideal answer (ability-to-transmit and occupancy truth matrices), and they should be robust to attempts to guess or otherwise game or subvert the intentions of the Challenge.

The final score, then, is a weighted sum of all of the alternative scores from Section 5.3.1. By choosing a zero-valued weight in the weight vector, a particular alternative score can be completely removed from consideration.

For the conditional occupancy, overestimated occupancy, and underestimated occupancy, we compute an average over the three user classes,

$$T_4 = (1/3)\left(T_4(P) + T_4(S) + T_4(U)\right), \tag{19}$$

$$T_5 = (1/3)\left(T_5(P) + T_5(S) + T_5(U)\right), \tag{20}$$

$$T_6 = (1/3)\left(T_6(P) + T_6(S) + T_6(U)\right). \tag{21}$$

Then the final score T_{final} is given by the weighted sum

$$T_{final} = \sum_{j=0}^{7} T_j \alpha(j), \tag{22}$$

where $\sum_{j=0}^{7} \alpha(j) = 1$. For relatively simple scene involving all three user classes, a Surrogate Participant System (SPS) produces ability-to-transmit matrices P and S and occupancy maps for the three user classes that yield the scores shown in Table 1.

α	T_{final}	Note
$\alpha = [1\ 1\ 1\ 1\ 1\ 1\ 1\ 1]/8$	0.907	Equal Weights
$\alpha = [2/9\ \ 2/9\ \ 2/9\ \ 1/15\ \ 1/15\ \ 1/15\ \ 1/15\ \ 1/15]$	0.933	Emphasize A-to-T
$\alpha = [1\ 1\ 1\ 0\ 0\ 0\ 0\ 0]/3$	0.962	Solely A-to-T
$\alpha = [1\ 0\ 0\ 0\ 0\ 0\ 0\ 0]$	0.979	Solely Basic A-to-T
$\alpha = [0\ 0\ 0\ 1\ 1\ 1\ 1\ 1]/5$	0.875	Solely Occupancy
$\alpha = [0\ 0\ 0\ 0\ 1\ 0\ 0\ 0]$	0.700	Solely Cond Occupancy
$\alpha = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 1]$	0.984	Solely Simple Occupancy

Table 1: Sample scores for Whisper-based processing of the Tranquility conceptual scenario. The score vector is $[0.98\ 0.92\ 0.99\ 0.99\ 0.70\ 0.83\ 0.87\ 0.98]$.

We intend to use the ”Emphasize A-to-T” weight vector in Table 1 to compute the final score.

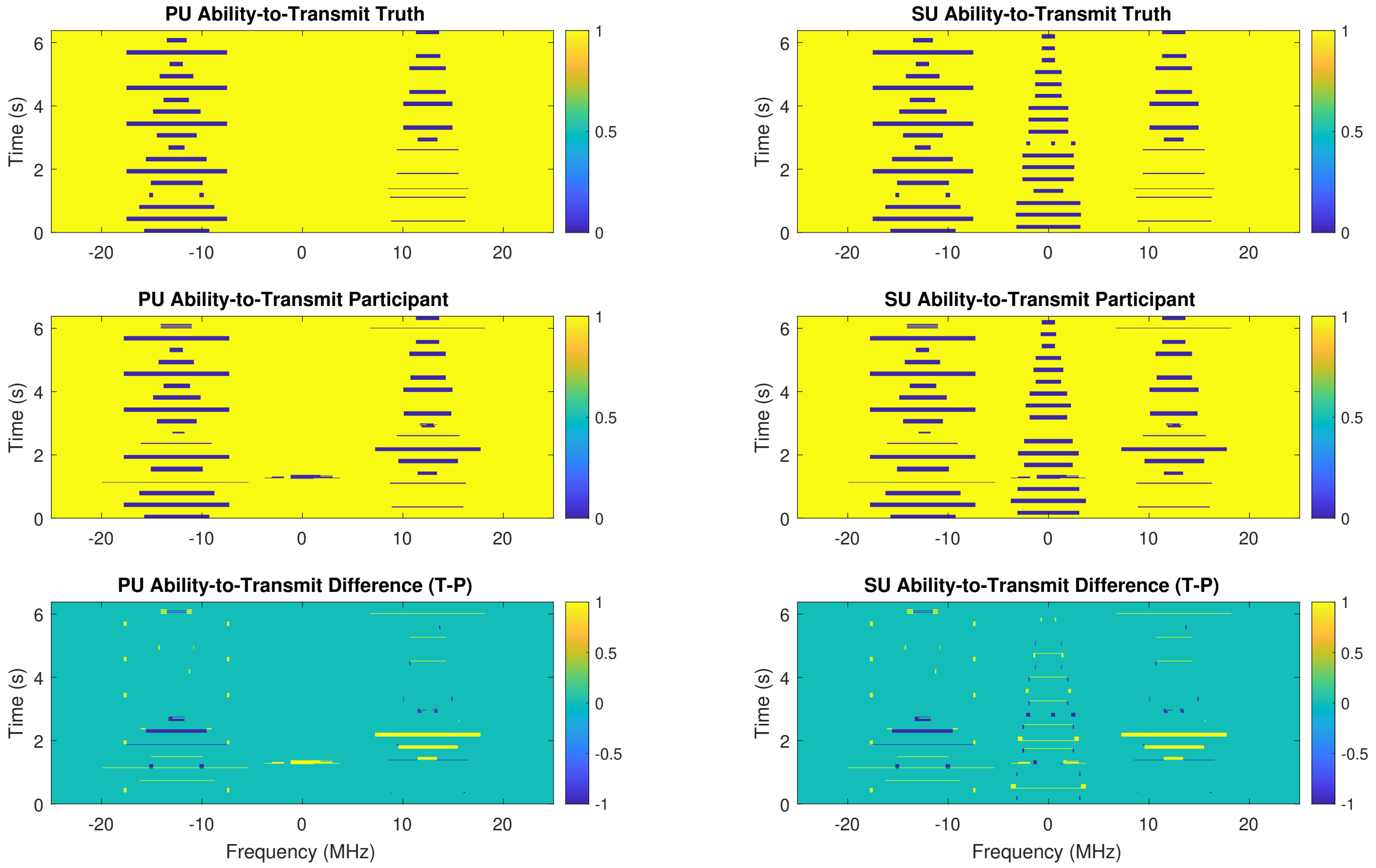

Figure 2: Ability-to-transmit results for SPS-based processing of a sample scene.

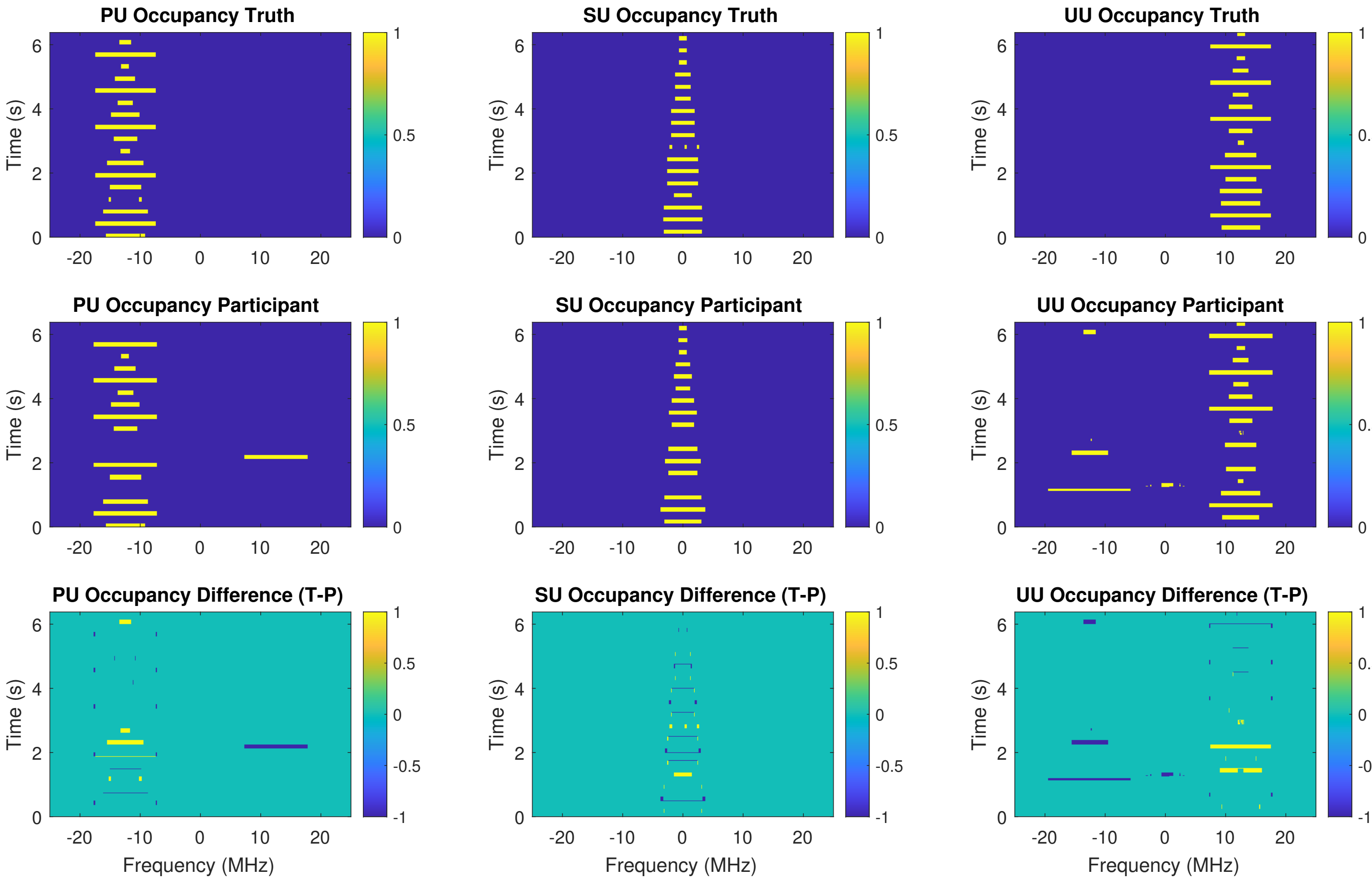

Figure 3: Spectral occupancy results for SPS-based processing of a sample scene.

6 Datasets Provided to Participants: Primary and Secondary Exemplars

Creating datasets that contain sufficient numbers of all involved primary, secondary, and urgent users, in all the combinations that might be seen in an RFSSC test scenario is a highly complex and time-consuming task. Moreover, such datasets may not be needed by signal-processing-oriented participants. On the other hand, to facilitate engagement with the Challenge, it may be desirable to provide *some* data that is representative of the Challenge test environments.

In no case will urgent-user datasets be created and disseminated to participants.

To balance the desire to supply some data to participants with the need to severely limit the labor cost of creating massive training/testing datasets, we employ a signal-definition approach that entails both mathematical descriptions of primary and secondary users' signals and the creation of *exemplar* datasets. An exemplar dataset will contain examples of the signal classes so that signals generated or simulated by participants can be checked against the exemplar to assess correctness. Exemplar datasets, while small, could be used in innovative ways to create sufficiently large training sets such that the Participant's inference engines can simultaneously produce good performance and good generalization. For example, existing data-augmentation techniques may be adapted or new ones devised so that the training sets greatly expand upon the provided exemplars, and yet are still grounded by the exemplar's correctness and applicability to the Organizer's sharing problem.

Good machine-learning system design always involves the generation and curation of high-quality datasets. In many cases this may entail generation from scratch using mathematical models, and it may also entail various forms of data augmentation to create sufficiently environmentally matched training and testing inputs. Therefore, we seek to split the responsibility of data generation between the Organizers and the Participants, which justifies the use of exemplars.

Because the RFSSC is more than conventional modulation recognition, where an algorithm or machine-learning system is devised to correctly produce a signal-type label for a single signal (nearly) centered at zero frequency, and because the provided exemplar signals are noise-free signals centered at exactly zero frequency, one or more *exemplar scenes* will also be provided along with the exemplar signals. This will make concrete the notion that the unknown-to-the-Participant RF scenes provided during competition will consist of multiple signals across a wide band of frequencies–the scene must be assessed, which means individual signals will likely have to be individually localized and processed before being passed to a neural network.

7 Data Rights

1. The US Government will retain ”rights in data” to all exemplar datasets provided to Participants. The USG will be free to use those datasets in future efforts of all kinds.
2. The US Government and a Participant organization will jointly retain ”rights in data” to all augmented datasets generated by the Participant. Both organizations will retain their ability to use the augmented datasets in perpetuity. The Participant will be required to place their augmented datasets on the Portal at the end of the Challenge.
3. The US Government and a Participant organization will jointly retain ”rights in data” to all non-augmented datasets generated by the Participant for the purpose of the RFSSC. Both organizations will retain their ability to use the augmented datasets in perpetuity. The Participant will be required to place their augmented datasets on the Portal at the end of the Challenge.

8 Data Files

8.1 Exemplar Signals

The signals to be used in the Challenge always correspond to a sampling rate of 50 MHz (complex-valued samples) and T seconds, for a total of $50T$ MSa. The notional format of the exemplar signal files is sigMF.

8.2 Exemplar Scenes

The exemplar to be used in the Challenge always correspond to a sampling rate of 50 MHz (complex-valued samples) and one second, for a total of 50 MSa. The notional format of the exemplar signal files is sigMF.

8.3 Challenge Scenes

The scenes to be analyzed in the competition phases of the Challenge always correspond to a sampling rate of 50 MHz (complex-valued samples) and ten seconds, for a total of 500 MSa. The notional format of the exemplar signal files is sigMF.

9 Mechanics of the Challenge

Round	Stage	Description	≈ Weeks	Supplied Data
Selection for Competition:				
1	1	Whitepaper preparation	5	Exemplar Scene 1 (50 MHz, 1 sec); Sharing Parameters
	2	Whitepaper scoring; Participant selection and awarding of prizes	3	
Total			8	
Preparation for Competition:				
2	1	Presentation of RFSSC		
	2	Provisioning of exemplars	1	Examples of all PU & SU signals
	3	Devise/train algorithms	4	
	4	Dry run with easy scenes	1	Exemplar Scenes 2–3; Sharing Parameters
Competition Round 1:				
3	1	Performers perform	4	Competition Scenes Set 1 (50 MHz, 10 sec); Sharing Parameters
	2	Organizers score submissions	2	
Competition Round 2:				
4	1	Performers perform	4	Competition Scenes Set 2 (50 MHz, 10 sec); Sharing Parameters
	2	Organizers score submissions	2	
Judging and Final Presentations:				
5	1	Final scoring and prizes	2	
	2	Winners brief funders/stakeholders	2	
Total			22	

Table 2: Basic RFSSC structure. The PU and SU example signals will consist of a small (relative to ML training) number of randomly generated examples for each distinct subclass.