The Problem of Sensing Unused Cellular Spectrum

Daniel Willkomm[1], Sridhar Machiraju[2,*], Jean Bolot[3], and Adam Wolisz[1,4]

[1] Telecommunication Networks Group, Technische Universität Berlin,
Einsteinufer 25, 10587 Berlin, Germany, `willkomm@tkn.tu-berlin.de`
[2] Google, Mountain View, CA 94043, USA, `machiraju@acm.org`
[3] Sprint, Burlingame, CA 94010, USA, `bolot@sprint.com`
[4] University of California, Berkeley, CA 94720, USA, `wolisz@ieee.org`

Abstract. Sensing mechanisms that estimate the occupancy of wireless spectrum are crucial to the success of approaches based on Dynamic Spectrum Access. In this paper, we present key insights into this problem by empirically investigating the design of sensing mechanisms applied to check the availability of excess capacity in CDMA voice networks. We focus on power-based sensing mechanisms since they are arguably the easiest and the most cost-effective. Our insights are developed using a unique dataset consisting of sensed power measurements in the band of a CDMA network operator as well as "ground-truth" information about primary users based on operator data. We find that although power at a single sensor is too noisy to help us accurately estimate unused capacity, there are well-defined signatures of call arrival and termination events. Using these signatures, we show that we can derive *lower bound* estimates of unused capacity that are both useful (non-zero) and conservative (never exceed the true value). We also use a combination of measurement data and analysis to deduce that multiple sensors are likely to be quite effective in eliminating the inaccuracies of single-sensor estimates.

Keywords: Cognitive radio, spectrum sensing, dynamic spectrum access.

1 Introduction

Dynamic Spectrum Access (DSA) is often viewed as a remedy against the spectrum scarcity caused by existing static spectrum allocation schemes. In DSA-based approaches primary users, who are often the licensed users of spectrum, have strictly higher priority than secondary users, who must vacate spectrum if and when it is needed by primary users.

A fundamental problem in DSA is: how can secondary users know whether or not primary users are using spectrum? This problem has been studied most

---

* The author was at Sprint when this work was done.

frequently in the case of secondary usage of spectrum licensed to TV broadcasters [10]. For TV bands as well as other usage scenarios, secondary users need to reach a binary decision, namely, whether a primary user is present or not. In this paper, we address spectrum licensed to another set of primary users – cellular telephony operators. Our focus on cellular bands is justified given the significant recent interest in implementing DSA in these bands [2–5, 8]. Such interest is fueled by the large number of devices and networks using cellular bands. In addition, as VoIP services and wireless data networks proliferate, cellular voice bands may see reduced loads. Secondary usage of such spectrum is an attractive way by which providers can extract value. We study CDMA cellular bands since they are one of the most widely-used cellular technologies (along with GSM) and we had access to the ground truth of such a network.

Since CDMA voice users share spectrum, the goal is not to detect whether a user is present or not but to identify the amount of *unused spectrum capacity*. Of course, even if there is unused spectrum capacity, secondary usage can have an impact on primary users. Hence, it is acceptable, perhaps even desirable, that some amount of spectrum is left un-utilized by secondary users. Thus, if secondary users want to use capacity X, we would like to check the availability of capacity $Y > X + \Delta$ beyond what is already used by primary users. Though understanding it is outside the scope of this paper, we do believe that – at the least – low-bandwidth applications (e.g., urban sensing) may be well placed to exploit unused capacity in CDMA voice networks.

In this paper, we investigate how secondary users can estimate unused spectrum capacity by utilizing sensing mechanisms. In particular, we focus on what are arguably the simplest sensing mechanisms – those that are based on single sensors considering power alone. Not only are these mechanisms simple but they are also cheap (thus enabling large-scale secondary user deployment). We phrase our investigation in terms of the following problem:

What information about primary user occupancy do spectrum measurements of power yield? Can this information be used to estimate the unused capacity that secondary users can exploit?

To conduct our investigation, we leverage a large and unique dataset containing **both** spectrum measurements of transmitted power as well as the corresponding "ground truth", i.e., the exact information about all the calls in progress at every point in time obtained from the call records collected at the network switches. *In other words, we capture both the actual behavior of the primary users (calls in progress, which can only be measured inside the network) as well as the estimated behavior of the primary users, as would be measured by secondary sensing users.* To our knowledge, this is the first study to execute and evaluate such synchronized measurements of sensing and ground-truth. Throughout this paper, we use the natural approach of estimating unused capacity by first estimating primary usage and subtracting it from the total capacity.

We start by exploring if sensed power can be easily converted to the amount of primary usage. We find that this way of estimating primary usage is not practical. Therefore, we focus on investigating other, indirect ways of estimating

unused capacity. In particular, we explore the possibility of identifying primary usage events (call arrivals or terminations). Using statistics of call durations that are published or well-known [7, 19], we can then convert the estimated intensity of event arrivals into estimates of primary usage. It turns out that sensed power, when averaged over time, contains distinct, well-defined power signatures corresponding to call arrival events and, to a lesser extent, call termination events. However, the noise in sensed power at any single time instant is large enough that we cannot exploit these signatures to accurately identify individual call events. We conclude that accurate estimation of primary usage (and hence unused capacity) using a single sensor is likely to be unachievable.

Power sensed by a single sensor can nevertheless be of value to secondary users since it can provide estimates of unused capacity that are conservative (underestimates) yet useful (not equal to zero). We show that a single sensor can provide such useful underestimates because of the following reason: even though our event identification algorithms are hampered by noise, they can be suitably configured so that they never overestimate unused capacity but still yield useful estimates of unused capacity. For example, in the moderately-loaded cells that we monitored in our experiments, we can correctly estimate that at least 5% of total capacity is almost always available for secondary use without ever overestimating unused capacity. Thus, we can use power measurements from a single sensor to at least support a low-bandwidth secondary application.

Finally, we show how accurate estimates of unused capacity can be obtained by using multiple sensors. Using our measurement-based characterization, we estimate the number of sensors required to achieve better accuracy.

2 Measurement Methodology

To evaluate if sensed power yields information about the underlying primary usage, we conducted a large number of sensing measurements spread over time and multiple locations. In addition, we simultaneously collected detailed information about primary usage. Thus, we were able to collect a large amount of unique data consisting of both measurements *and* ground truth. Our measurements were collected in the band used by a CDMA-based cellular operator, which is an important target for dynamic spectrum access as described in the introduction. In total, our experiments yielded approximately 90 GB of ground truth data on primary usage and 14 GB of data with spectrum power measurements. In this section, we describe our measurement methodology in detail and also the salient aspects of the collected datasets.

2.1 Ground Truth on Primary Usage

To compute the ground-truth, we use call records collected at the switches of a CDMA-based cellular operator. These records capture the start time, duration, initial and final sector as well as assigned carrier[5] of the voice calls made in the

---

[5] Each cell sector may be assigned one or more carriers – a 1.25 MHz portion of the spectrum. Each carrier is capable of supporting tens of calls.
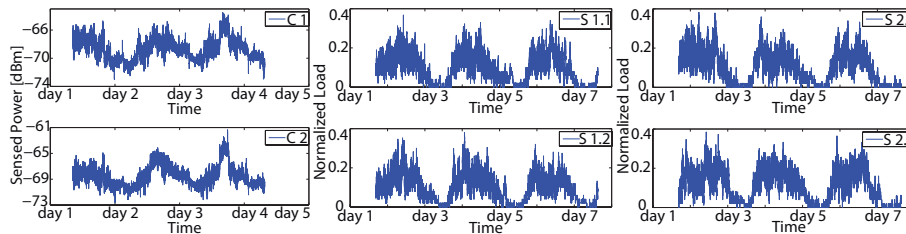
Fig. 1. Dataset $D1$: **(a)** Sensed power over time for carriers $C1$ and $C2$. **(b)** Normalized load for strongest cell sector $S1$. **(c)** Normalized load for 2nd strongest cell sector $S2$.

network. The call duration reflects the RF emission time of the call and captures precisely what we want – the duration of primary usage. All timestamps were measured with a resolution of a few milliseconds.

Using the call records, we calculate the ground-truth usage in each carrier of a cell sector. We split the call records based on the sector and carrier. We create two records for each call corresponding to its initiation and termination. Then, we sort these records in order of their time to get a sequence of call *events* that are known to have occurred in each sector/carrier. Using the sorted list of events, we also calculate the *load* of each sector and carrier. We do so by maintaining a running count of the number of ongoing calls. This count is increased by $+1$ when a call begins and decreased by -1 when a call terminates.

Since the switches only record the initial and final sector of each call, we are unable to account for the spectrum usage in other sectors that the user may have visited in between. This implies that our list of events is accurate but not necessarily complete (it may not count call initiation and termination events caused by handover). Similarly, for calculating load, we assign the whole call to the initial sector/carrier. We also try other approximations of load by, for example, assigning the first half of the call to the initial sector/carrier and the last half to the final sector/carrier. Since these approximations do not alter our results significantly, we do not provide them. Furthermore, we believe that full mobility information is unlikely to change our results either.

2.2 Power Sensing

We collected our power measurements using a W1314A Multi-band Wireless Measurement Receiver from Agilent [1]. We used the Model 110 so that we can sense the uplink and downlink CDMA bands (in the 1900 MHz frequency range) used by the network whose call records we had access to. This receiver captured in real-time power measurements from all 12 1.25 MHz carriers belonging to a single provider. The power measurements were reported twice or thrice a second on average. Unless stated otherwise, we convert this raw data into per-second averages (computed using 2 or 3 readings). Our wireless measurement receiver is a sophisticated piece of equipment and can be viewed as being capable of the *most accurate* power based sensing that a secondary user can perform.

We collected multiple datasets at four different urban locations ($L1$ to $L4$). For each of the locations the power measurements were collected over multiple

days and spanned all possible 1.25 MHz carriers of the CDMA band. We use $C1$, $C2$, ... to refer to the various carriers. $L1$ was within line-of-sight and about 0.5 miles from the nearest base station. $L2$ was at a similar distance but not with line-of-sight of the closest base station. In both $L1$ and $L2$ the antenna was placed close to a window at the second floor. $L3$ and $L4$ were about twice the distance (1 mile) to the closest base station, on the ground floor and heavily shadowed from the closest base station.

Our analysis showed the following trends: datasets collected at the same location (at different times) showed similar results. Furthermore, the datasets collected at locations $L1$ and $L2$ illustrated better results than those collected at locations $L3$ and $L4$. This was due to the heavy shadowing at the latter locations. On account of this and space limitations, we provide results only for $L1$ and $L2$ in this paper. In particular, we use two datasets, which we label $D1$ and $D2$. $D1$ was collected over a period of 3 days at location $L1$ and $D2$ was collected over a period of 4 days at location $L2$. In this paper, we present results for two carriers $C1$ and $C2$, which were the most active carriers in the two locations.

For each experiment, we also recorded the identities of the cell sectors with the strongest pilots. For each dataset, we refer to the cell sectors as $S1$, $S2$, ... in decreasing order of pilot strength. We observe that, for both of our datasets, this order reflected the distance of the sectors from the measurement locations. Usually a cell sector has activity in more than one carrier. We denote this by referring to the activity of sector 1 in carrier 1 with $S1.1$, in carrier 2 with $S1.2$, etc. Note that, for each of these cell sectors, we computed ground-truth (events and load) using the call records as described in Sect. 2.1.

3 Estimation via Power Thresholds

In this section, we present how well power is correlated with primary usage. We obtain these results by investigating the following question: Can a simple scheme based on mapping sensed power information to load information work? We start with some preliminary observations and data analysis for synchronizing the power measurements and ground truth.

3.1 Power-Load Correlation

In Fig. 1, we plot the power sensed and network ground truth for dataset D1 (downlink). We plot results for the two strongest sectors $S1$ and $S2$ and two carriers $C1$ and $C2$ used by both sectors. For data confidentiality reasons, we normalize the load values by a randomly chosen number so that the absolute load is obfuscated while preserving the trends. Notice that the day/night variations of the load are clearly visible. This is not surprising since it corresponds to levels of human activity and has also been observed in prior work [9, 19]. The plots showing the sensed power also illustrate a distinct diurnal pattern with higher power levels during the day and lower levels at night.
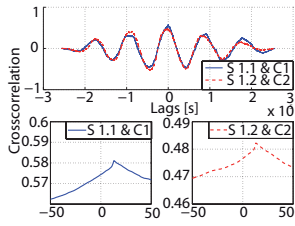
Fig. 2. Dataset $D1$: Cross-correlation of sensed power and ground truth load.
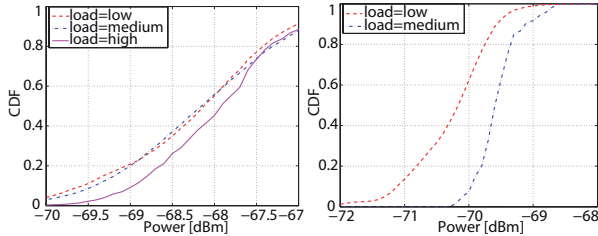
Fig. 3. Dataset $D1$: Distribution of power levels when the load varies. **(a)** Daytime **(b)** Nighttime. We divide the observed load levels into three levels.

Cross-correlation is a simple and well-known metric that we can use to measure the extent to which power "tracks" load. We plot the cross-correlation between sensed power and load for several lags in Fig. 2. The maximum is not reached at lag 0 since the clocks used for collecting call records and power measurements are only synchronized to within a few seconds of each other. In addition, for each sector/carrier, we see peaks separated by roughly one day. These local maxima are caused by the underlying diurnal pattern of the load. Overall, it is clear that the sensed power is indeed well correlated with load.

It turns out that the cross-correlation curves for all sector/carriers reach a maximum at the same lag of 13 seconds as shown in Fig. 2. We observed maxima at a similar lag with dataset $D2$ as well. For the rest of this paper, we, thus, use the above lag to synchronize the power measurements and call records.

3.2 Naive Threshold-based Scheme

We start investigating if there is a unique mapping of power level to sector load with Fig. 3**(a)**. We plot the distribution of the sensed power levels for various coarse-grained levels of load during peak hours of the day (12PM to 6PM). As expected, with increasing load, the power levels tend to increase. But, observe that for different values of load, the same power levels are often seen. Moreover, when the load is low, the power levels are more spread out. This indicates that it is challenging to distinguish small changes in load using static power alone.

Power levels are better able to separate the coarse-grained measures of load at night (12AM to 6AM) as shown in Fig. 3**(b)**. Note that, since the load at night is never high, we only show two levels. These levels are relatively well separated: for example, the power level is below -70 dBm for 60% of the time when the load is low as opposed to 5% of the time we have medium load. Such power-based thresholds separating coarse-grained load levels such as low, medium and high are likely to depend on location. For example, in separate short experiments, we found that the average power levels close (about 100 m) to a base station were around -50 dBm. Though we did not find much variation at a given location within a few days, it is unclear how long thresholds at a location remain valid.

To summarize, load estimation based on power thresholds (that are gleaned on a per-location basis) can provide coarse-grained information about load especially at night. Though fine-grained information is difficult to extract, such
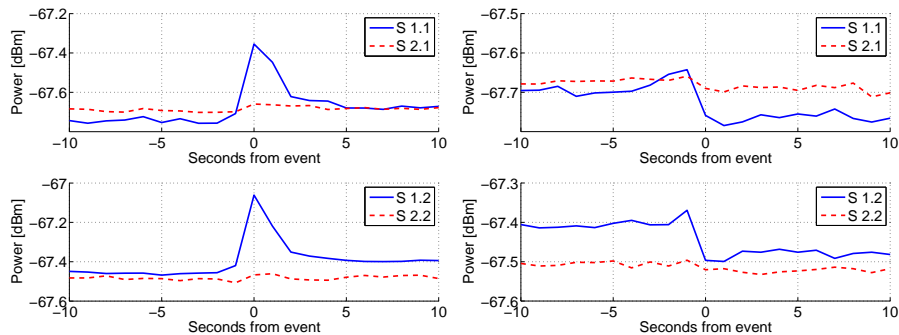
Fig. 4. Dataset $D1$: 10-second *t-average* plots for (downlink) power for $C1$ (top) and $C2$ (bottom) during **(a)** Call initiation events, and, **(b)** Call termination events.
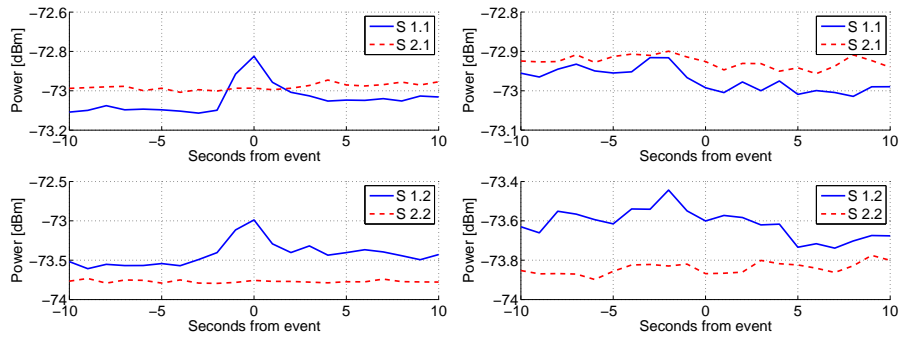

Fig. 5. Dataset $D2$: 10-second *t-average* plots for (downlink) power for $C1$ (top) and $C2$ (bottom) during **(a)** Call initiation events, and, **(b)** Call termination events.

coarse-grained information might be sufficient to decide whether or not to start secondary usage. The above threshold-based scheme is a *black-box* estimation of load in the sense that no information is required from the cellular operator.

4 Event Signatures

During our analysis, we found that sensed power contains information about the "first derivative", i.e., change of primary usage load. These changes are typically due to call initiation and termination in a CDMA voice network.

The sensed power in our datasets shows often jumps and drops when events occur. To understand if such event *signatures* exist and characterize them, we examine all initiation and termination events spread across multiple days in our datasets. Since the average power across these days may not be stationary we rely on what we refer to as *t-average* plots: we extract the power for a short time periods before *and* after events and average the sensed power across all initiation and termination events. This allows us to look at time-averaged characteristics of power when events occur even in the absence of stationarity – a key advantage.

In Fig. 4, we show *t-average* plots capturing the average behavior for 10 seconds before and after call initiation and termination events of dataset $D1$.

In this figure, we are considering only the downlink bands. We notice two key patterns for the events involving the strongest cell sector $S1$:

– On average, call initiation events are characterized by a spike that is about $0.3 - 0.4$ dBm followed by a general increase of power afterwards that is around $0.05 - 0.1$ dBm. This trend is clearly seen during the daytime and nighttime as well (not shown). We believe that this spike is easily explained by the CDMA downlink power control loop [18]. This loop ensures that, when a call starts, the base station transmits with high power. The power level is then reduced to a minimal level while maintaining call quality using the rapid closed-loop power control of CDMA. The increase of $0.05 - 0.1$ dBm reflects the increased power due to a new call, of course.
– On average, call termination events are characterized by a dip of at least $0.05 - 0.1$ dBm (a bit higher during nighttime) immediately after the event. This reflects our intuition that the call termination corresponds to lesser power being emitted by the base station.

Thus, there exist well-defined power signatures for initiation and termination events in CDMA networks. Figure 4 also shows that there are no visible signatures corresponding to the events of the second strongest cell sector $S2$.

Recall from Sect. 2.2 that we collected a second dataset $D2$ from experiments conducted at another location. To verify that the event signatures persist across locations, we plot *t-average* plots for $D2$ in Fig. 5. The call initiation signature corresponding to events of the strongest sector $S1$ continues to be clearly seen for both carriers. As before, there is no signature corresponding to events of the second strongest sector $S2$. Surprisingly, the call termination signature for $D2$ is less clear. The location at which dataset $D2$ was collected had no line-of-sight to $S1$ or $S2$ which is the likely reason behind the weak termination signature.

We also use *t-average* plots (not shown due to lack of space) to investigate if such signatures are also present in the sensed power of uplink bands. We find no identifiable signatures corresponding to initiation or termination events for both $S1$ and $S2$. The absence of signatures in uplink power is not surprising given that the average power levels are about $25 - 30$ dBm lower than the downlink power measurements. Such lower power levels are likely due to the stricter power budget of end-user devices. Also, the sources of uplink power are end-user devices, which are spatially distributed. We expect to see signatures when such devices are nearby. To verify this hypothesis, we conducted "active" experiments by initiating phone calls using a mobile handset located near our power sensor. When these calls were initiated, we did observe identifiable spikes in power similar to the downlink initiation signature. These experiments confirm that uplink sensing is of use only if the sensor is close to all end-user devices. Since this is physically impossible, we do not further investigate uplink sensing in this paper.

5   Event Detection

In the previous section, we found the presence of well-defined signatures corresponding to call initiation and termination events on the downlink CDMA

channels. In this section, we show that such average-case signatures do, however, not translate into algorithms for accurate event detection.

5.1 Discriminators of Initiation Signature

The *t-average* plots discussed in the previous section indicate that call initiation can potentially be identified by detecting *spikes* in the sensed power. Referring to the *t-average* plots of Fig. 4, we see that there are roughly three time periods: during, before and after call initiation. The spike occurs during the call initiation and is significantly higher than the power before and after. We also expect the power after call initiation to be higher than before.

Based on the above discussion, we are motivated to consider three intuitive discriminators of initiation signatures. We use $P(\cdot)$ to denote the power (as a function of time) and T to represent the second when a call is initiated. We divide a contiguous period of time around T into three periods: the first period of call initiation consisting of a window of w seconds including and after T, a window of w_- seconds prior to this period, and, a window of w_+ seconds after the call initiation period. We calculate the average power in each of these three periods and define the criteria for the 3 discriminators as follows:

1. The difference between the power in the call initiation period and the power in the period before call initiation is larger than a threshold τ_1, i.e.,

$$\overline{P([T, T+w-1])} - \overline{P([T-w_-, T-1])} \geq \tau_1.$$

2. The difference between the power during the call initiation period and the power thereafter is larger than a threshold τ_2.

$$\overline{P([T, T+w-1])} - \overline{P([T+w, T+w+w_+-1])} \geq \tau_2.$$

3. The difference between the power after and before call initiation is larger than a threshold τ_3.

$$\overline{P([T+w, T+w+w_+-1])} - \overline{P([T-w_-, T-1])} \geq \tau_3.$$

An advantage of using a window of several seconds in each period might be the potential reduction in the variance of the estimated power. At the same time, larger periods may be polluted by other call initiation and termination events.

To better understand the impact of the various parameters including the window sizes and thresholds, we rely on so-called *cCDF* (*complementary CDF*) plots. Consider the first discriminator. Recall that it looks at the difference during and before call initiation and expects this difference to exceed a threshold when a call is initiated. A *cCDF* plot shows if this discriminator is justified by plotting the distribution of the difference for all T when a call was initiated, and, compares it with the distribution for all T when a call was not initiated. Since we are interested in the number of cases that the difference exceeds a threshold, we plot the *cCDF* (the CDF subtracted from 1) and experiment with several choices of w and w_- (plots not shown due to lack of space).
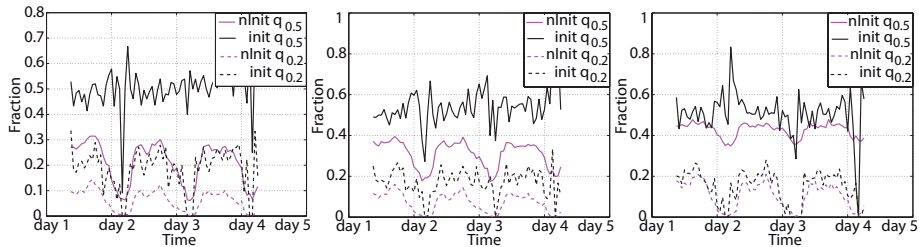
Fig. 6. Dataset $D1$: Fraction of initiation events and non-events that satisfy the discriminators defined by **(a)** Criterion 1. **(b)** Criterion 2. **(c)** Criterion 3.

The higher we choose τ_1 to be, the fewer call initiation events satisfy the criterion. However, as we make τ_1 smaller, more seconds during which no call was initiated satisfy the criterion. The sweet spot appears to be at around $0.4 - 0.5\,\text{dBm}$, below which the latter increase faster than the former. We also get marginally better results when we define the period during call initiation as consisting of exactly 1 second. Larger windows do not improve our results. It seems that the power during additional seconds is not as high as that of the first second. Hence, any potential variance reduction from the additional power measurements comes at the cost of eliminating the signature itself. The size w_- of the period before call initiation impacts the results to a lesser extent, leading to marginally better results with a window of size 4.

Using the *cCDF* plots for the other two discriminators we find that they exhibit similar behavior, namely, smaller windows are better. $0.3 - 0.5\,\text{dBm}$ appears to be a good threshold value of τ_2. However, the third criterion is not as useful since the difference in power before and after call initiation is not as clear.

The success of our criteria may vary with time which is explored in Fig. 6. We choose the thresholds corresponding to the 0.5 and 0.2 *cCDF* quantiles. We then apply the criteria with the respective thresholds on an hourly basis and plot the fraction of initiation events and "non-events" satisfying them in Fig. 6.

As expected, the fraction of initiation events satisfying each criterion does not vary significantly and stays around the quantile value (0.5 or 0.2) used to choose the thresholds. However, the fraction of "non-events" passing the criteria show a clear dip during night time for both thresholds. This implies significantly better performance during night time on account of lesser noise in power.

5.2 Initiation Detectors

We now investigate how the various discriminators and their criteria can be used for the best possible detection of initiation events. Due to lack of space, we do not focus on detecting call terminations since discriminators based on call termination events do not perform as well as those based on call initiation.

We quantify discriminator performance, by looking at the *detection probability*, i.e., the probability of detecting an initiation event given that a call was really initiated. In Fig. 7**(a)**, we plot the detection probability when each of our three discriminators are used on dataset $D1$. For each detector, we vary the
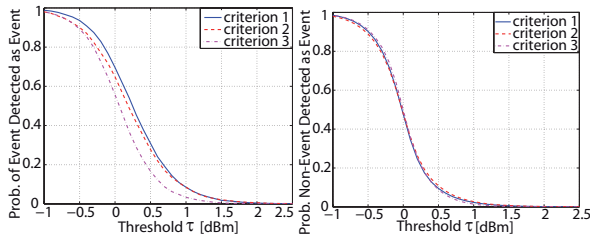
Fig. 7. Dataset $D1$: **(a)** Fraction of initiation events identified by the detectors **(b)** Fraction of non events declared as initiation events by the detectors based on the 3 criteria with varying threshold values.
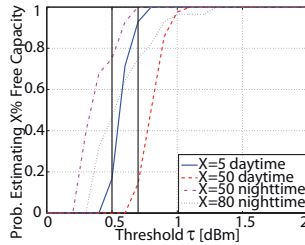
Fig. 8. Dataset $D1$: Probability of estimating at least $X\%$ of unused capacity.

corresponding threshold from -1 to $2.5\,\mathrm{dBm}$. As expected, the detection probability reduces with increasing thresholds. In Fig. 7**(b)**, we show the probability of detecting a call initiation event, although no call was initiated.

Comparing Fig. 7**(a)** and Fig. 7**(b)** shows a common problem using energy detection: Choosing a threshold τ_1 that achieves a high detection probability results in many non-initiation events to be declared as initiation event and vice versa. E.g., using criterion 1, a threshold of $\tau_1 = 0\,\mathrm{dBm}$ results in detecting close to 75% of the initiation events, but also in mistakenly detecting about 50% of the non-events as initiations. In contrast, a threshold of $\tau_1 = 0.5\,\mathrm{dBm}$ detects only 10% of the non-events as initiations but fails to detect 70% of the initiations.

6 Estimating Unused Capacity

In this section, we discuss how we can use power sensed by a single sensor to derive useful estimates of unused capacity. Specifically, we find lower bounds for the unused capacity so that – at least – low-bandwidth applications can utilize it. Note that these estimates are useful because they are significantly greater than zero. We compute lower bounds using the call initiation detectors to estimate call arrival rates. We then estimate the load in the system using Little's law as $E[k] = \lambda E[b]$ where $E[b]$ is the mean call duration and λ the estimated arrival rate. We do assume that we have **partial** information about the system being studied, namely, the average call durations. Such information can be obtained from previous studies [7, 19] (since mean call durations are quite stable over time) or directly from providers. Since we use such information, we refer to it as a *gray-box* approach (in-between *black-box* and *white-box* approaches).

In Fig. 8 we show how well we can achieve our goal: finding unused spectrum to satisfy the secondaries bandwidth requirement. We show results for dataset $D1$ and criterion 1. We divide our dataset into hour-long time periods and calculate average values within these time periods. Remember, that it is crucial to never overestimate the unused capacity. Using the whole dataset, the maximum threshold would be $\tau \leq 0.5\,\mathrm{dBm}$ resulting in $X = 5\%$ of capacity unused in more than half of the time periods (figure not shown). If we use different thresholds for daytime (9am to 10pm) time periods and nighttime (11pm to 8am) time periods,

we can improve performance significantly. The maximum thresholds so that we never overestimate are $\tau \leq 0.7\,\text{dBm}$ and $\tau \leq 0.5\,\text{dBm}$ respectively. These are indicated by the solid black lines in Fig. 8. This figure also shows that, using the maximum daytime threshold ($\tau = 0.7\,\text{dBm}$), we correctly estimate 5% (50%) of total capacity is unused almost always (in 14% of the daytime time periods). For the maximum nighttime threshold ($\tau = 0.5\,\text{dBm}$) we correctly estimate that 50% (80%) of total capacity is unused in 75% (46%) of the nighttime hours.

For dataset $D2$, the results are similar though the maximum thresholds (so that we never overestimate) are closer to $1\,\text{dBm}$. These results show that power measurements at a single sensor can be quite useful especially for secondary applications such as urban sensing, which have low bandwidth demands but strict power constraints. Given the differences we observed between $D1$ and $D2$, local calibration might be necessary to choose the appropriate power thresholds.

7   Towards Spatial Diversity

An alternative approach is to improve the accuracy of event detection. Clearly, this is hard to achieve with a single sensor because the signatures of individual events are obfuscated by additive white noise, which makes accurate detection very hard especially in the low SNR regime [15, 17]. The natural way to improve event detection accuracy would be to use multiple sensors in spatially diverse locations so that we can eliminate the white noise. There are various proposals for cooperative spectrum sensing approaches in the literature, e.g., [13, 14, 16].

We now use our dataset to quantify the benefits of cooperative sensing. In particular, we consider the approach of soft decision combining as described in [13] and make a natural assumption of zero-mean additive Gaussian noise with variance σ^2. With k distributed (and independent) sensors, the white noise in the average sensed power can be approximated as a zero-mean normal variable $\mathcal{N}(0, \frac{\sigma^2}{k})$ with variance $\frac{\sigma^2}{k}$ by the Central Limit Theorem.

We calculate the quality of event detection under the above model: Assume that a call is initiated at time T and the first criterion of Sect. 5 with $w = w_- = 1$. At time T, the criterion shows a power spike of about $0.4\,\text{dBm}$ plus the difference between the white noise at T and $T - 1$. Assuming that white noise at these time instants is independent, their difference is a zero-mean normal variable with variance $\frac{2\sigma^2}{k}$. Hence, the criterion will have a false negative with probability P_{fn} and false positive with probability P_{fp}

$$P_{\text{fn}} = P\left(\mathcal{N}\left(0, \frac{2\sigma^2}{k}\right) + 0.4 \leq \tau\right) \qquad P_{\text{fp}} = P\left(\mathcal{N}\left(0, \frac{2\sigma^2}{k}\right) \geq \tau\right) ,$$

where τ is the decision threshold. Given a maximum tolerable false positive and false negative probability, we can solve the above to estimate the minimum number of sensors (k) required. For example, for $P_{\text{fn}} \leq p$ and $P_{\text{fp}} \leq n$ we get:

$$\sqrt{k} \geq \max\left(\left|\frac{q_{1-p}\sqrt{2}\sigma}{\tau_1}\right|, \left|\frac{q_n\sqrt{2}\sigma}{\tau_1 - 0.4}\right|\right)$$

Here, q_y is the value at which the quantile of the standard normal distribution is y. We estimate σ using $D1$. Since we do not want to capture temporal mean variations, we remove the moving average of the previous 5 seconds for each power reading. This yields an estimate $\sigma = 0.39$ with which we can achieve less than 10% false negatives and false positives using $k = 13$ sensors and $\tau = 0.2\,\mathrm{dBm}$. In the previous section, we saw that it may be desirable to have a much smaller fraction of false positives than false negatives. It turns out that we can achieve at most 1% false positives and 10% false negatives with $k = 26$ sensors and $\tau = 0.25\,\mathrm{dBm}$. Deploying such numbers of sensors per cell could very well be economical, especially if existing consumer devices can be leveraged.

8   Related Work

In recent years, a lot of measurement studies [6, 11, 12] have been carried out to show the under-utilization of the licensed spectrum. Though these studies show the abundance of temporally unused spectrum, they give little insight into the dynamic behavior of the licensed users legally operating in those bands.

The authors of [9] estimate the load in the New York cellular bands (CDMA as well as GSM) based on spectrum measurements. However, in addition to pure power measurements, the CDMA signals are demodulated to determine the number of active Walsh codes (i.e., the number of ongoing calls). To determine the number of calls in the GSM bands, image processing of the spectrogram snapshots is used. This is in contrast to our study, which is based on power measurements and uses minimal processing. In addition, to our best knowledge, there is no study which correlates spectrum measurements with the actual load as recorded by the system. Sensing for TV bands has been previously studied, for example in [10]. They found that energy detection with multiple sensors is often better than feature detection. However, their results are for the relatively static TV bands and not for cellular bands.

9   Conclusions

We used a unique set of simultaneous sensing and network measurements to study the problem of sensing-based estimation of unused capacity in cellular spectrum. When averaged over time, we found well-defined signatures of call initiation and termination events using the power at a single sensor. However, sensing noise makes it challenging to use these signatures to estimate unused capacity by identifying call events. We found that useful underestimates can nevertheless be computed especially for low-bandwidth secondary applications. Alternatively, we can obtain accurate estimates by using multiple sensors. To our knowledge, our work is the first detailed study of how well sensing works in CDMA networks, which are often viewed as candidates for DSA. In the future, we intend to design and evaluate better sensing algorithms including those based on multiple sensors.

Bibliography

[1] Agilent: W1314a datasheet. `www.agilent.com`, [online]

[2] Al Daoud, A., Alanyali, M., Starobinski, D.: Secondary pricing of spectrum in cellular CDMA networks. In: Proc. of IEEE DySPAN 07 (2007)

[3] Alyfantis, G., Marias, G., Hadjiefthymiades, S., Merakos, L.: Non-cooperative dynamic spectrum access for cdma networks. In: Proc. of IEEE GLOBECOM 07 (2007)

[4] Buddhikot, M., Ryan, K.: Spectrum management in coordinated dynamic spectrum access based cellular networks. In: Proc. of IEEE DySPAN 05 (2005)

[5] Chen, D., Yin, S., Zhang, Q., Liu, M., Li, S.: Mining spectrum usage data: a large-scale spectrum measurement study. pp. 13–24 (2009)

[6] Chiang, R., Rowe, G., Sowerby, K.: A quantitative analysis of spectral occupancy measurements for cognitive radio. In: Proc. of VTC Spring 07 (2007)

[7] Guo, J., Liu, F., Zhu, Z.: Estimate the call duration distribution parameters in GSM system based on k-l divergence method. In: Proc. of WiCom 07 (2007)

[8] Hamouda, S., Hamdaoui, B.: Dynamic spectrum access in heterogeneous networks: Hsdpa and wimax. In: Proc of IWCMC 09 (2009)

[9] Kamakaris, T., Buddhikot, M., Iyer, R.: A case for coordinated dynamic spectrum access in cellular networks. In: Proc. of IEEE DySPAN 05 (2005)

[10] Kim, H., Shin, K.G.: In-band Spectrum Sensing in Cognitive Radio Networks: Energy Detection or Feature Detection? In: Proc. of ACM Mobicom 2008 (2008)

[11] MacDonald, J.T.: A survey of spectrum occupancy in chicago. Tech. rep., Illinois Institute of Technology (2007)

[12] McHenry, M.A., Tenhula, P.A., McCloskey, D., Roberson, D.A., Hood, C.S.: Chicago spectrum occupancy measurements & analysis and a long-term studies proposal. In: Proc. of TAPAS 06 (2006)

[13] Mishra, S.M., Sahai, A., Brodersen, R.W.: Cooperative sensing among cognitive radios. In: Proc. of IEEE ICC 06 (2006)

[14] Pham, H.N., Zhang, Y., Engelstad, P.E., Skeie, T., Eliassen, F.: Optimal cooperative spectrum sensing in cognitive sensor networks. In: Proc. of IWCMC 2009 (2009)

[15] Sahai, A., Tandra, R., Mishra, S.M., Hoven, N.: Fundamental design tradeoffs in cognitive radio systems. In: Proc. of TAPAS 06 (2006)

[16] Sun, C., Zhang, W., Letaief, K.B.: Cooperative spectrum sensing for cognitive radios under bandwidth constraints. In: Proc IEEE WCNC 07 (2007)

[17] Tandra, R., Sahai, A.: Snr walls for signal detection. IEEE J. Sel. Topics Signal Process. 2(1), 4–17 (Feb 2008)

[18] Vanghi, V., Damnjanovic, A., Vojcic, B.: The cdma2000 System for Mobile Communications: 3G Wireless Evolution. Prentice Hall PTR (2004)

[19] Willkomm, D., Machiraju, S., Bolot, J., Wolisz, A.: Primary Users in Cellular Networks: A Large-scale Measurement Study. In: Proc. of IEEE DySPAN 08 (2008)