Spectrum Data Mining: Measurement-Driven Insights for Sustainable Spectrum Management

Amir Ghasemi
Communications Research Centre Canada
Email: firstname.lastname@canada.ca

***Abstract*—Communications Research Centre Canada (CRC) has deployed a network of spectrum sensors with high time and frequency resolution for continuous monitoring and analysis of spectrum activity in different bands. In the present work, the radio frequency (RF) measurement data collected over a long period of time is used to characterize the spectrum usage of incumbent users in Land Mobile Radio (LMR) bands. Statistical analysis of measured data indicates daily and weekly spectrum usage patterns, especially in the channels used by public safety users. The identified activity patterns are then leveraged to predict future spectrum usage, thereby enabling a data-driven proactive approach to spectrum assignment and management, where spectrum shortage and oversupply across different networks can be predicted and managed. We further investigate correlations between spectrum activity and external factors such as severe weather and special events. The findings confirm existence of such correlation and provide further insights for dynamic spectrum assignment in Land Mobile Radio bands.**

I. Introduction

Land Mobile Radio (LMR) systems are often used by government agencies, municipal services, and a variety of commercial users to enable (mostly push-to-talk) voice communications among a group of users. These systems are critical, especially in public safety and emergency response operations, where multi-cast voice communications (e.g., between first responders and a dispatch centre) are needed.

Typically, frequency channels in the LMR bands are assigned to different licensees based on the information received from the users when they apply for a licence (e.g., transmitter location, antenna height, emitted power, etc.). In addition, regulators primarily have to rely on predefined models for the (voice) traffic load and propagation loss to estimate the impact of a potential assignment on other users. In the absence of measurement data showing where, when, and how different parts of the spectrum are being used, regulators have to be more cautious in reusing the spectrum in order to avoid interference among different radio systems.

To address the knowledge gap described above, Communications Research Centre Canada (CRC) has developed a prototype spectrum monitoring system which, using multiple sensors distributed over a geographical area, continuously monitors the spectrum usage in selected frequency bands below 6 GHz, and uploads the collected measurement data to the cloud for further processing and analysis [1]. Since its deployment in Ottawa, Canada in February 2016, this system has gathered terabytes of data about spectrum usage at different locations, times, and frequencies, thereby enabling data-driven approaches to spectrum assignment, sharing, and interference resolution [2].

The work presented herein describes preliminary research toward a better understanding of spectrum usage, and some of the underlying factors driving it, by exploiting historical RF data collected by CRC's spectrum monitoring system. The exploratory analysis is focused on a subset of these measurements in the 800 MHz LMR band (specifically the public safety subband at 866-869 MHz). As we will illustrate later, LMR spectrum usage is non-uniform across different spectral ranges and exhibits time and frequency patterns that could be exploited to improve users' timely access to the spectrum. This paper's focus is on temporal characterization of these bands and exploring the existence of correlations between intensity of spectrum usage and external factors, such as special events or severe weather.

The remainder of this paper is organized as follows. Section II describes the spectrum measurement setup for the bands of interest in this study. Section III discusses the spectrum occupancy analysis and prediction, as well as some external factors affecting the spectrum usage. Finally, conclusions are presented in Section IV.

II. Spectrum Measurement System

CRC has developed a spectrum monitoring system designed to provide detailed spectrum usage information for research purposes. The prototype includes a wide range of both fixed and mobile sensors covering different frequency ranges to provide a continuous and wideband monitoring capability below 6 GHz [1]. The system's sensors are networked and their near real-time RF measurements are reported to a cloud-based processing entity for subsequent retrieval and analysis work.

As the focus of the present work is on the time and frequency dimensions, spectrum data used for the analysis was collected from a single, fixed monitoring station. The exact measurement location is shown in Fig. 1 (denoted by the red marker). Operationally, this sensor is a FFT-based receiver covering 138 MHz to 952 MHz

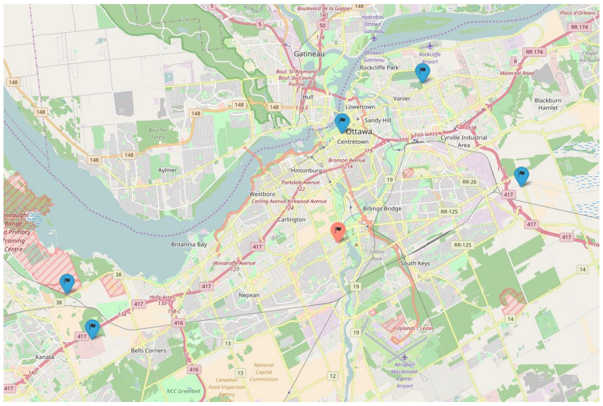


Fig. 1. Location of fixed measurement sites covering the LMR bands

(a) (b)

Fig. 2. View of (a) line-of-sight from measurement antenna towards downtown and (b) spectrum monitoring station equipment.

with a fast digital sweep and an instantaneous bandwidth of 24.76 MHz. For the bands of interest, it uses an omnidirectional and vertically polarized wideband discone antenna installed at a height of 80 m above ground on the roof of a building approximately 5 km south of downtown Ottawa.

Fig. 2 shows the view from the antenna towards downtown Ottawa. For the measurements used in this study, the average band sweep time was approximately 0.3s with a resolution bandwidth of 1.984 kHz. Such a sweep time provides sufficient time resolution to capture a variety of LMR traffic patterns. For each measured frequency channel, the detection power threshold was set to a margin over the estimated noise floor to account for temporal noise variations. The threshold was set using the method in [3] based on deeming a probability of false alarm of 10^{-5} to be the maximum acceptable error.

Fig. 3 shows a 5-minute snapshot of high-resolution spectrum measurements for downlink public safety LMR band (866-869 MHz) in Ottawa. Most radio systems operating in this band are trunking systems providing voice communications to a variety of public safety and municipal users. The continuously active channels in Fig. 3 (i.e. the uninterrupted horizontal lines) typically represent the control channels of trunked radio systems used

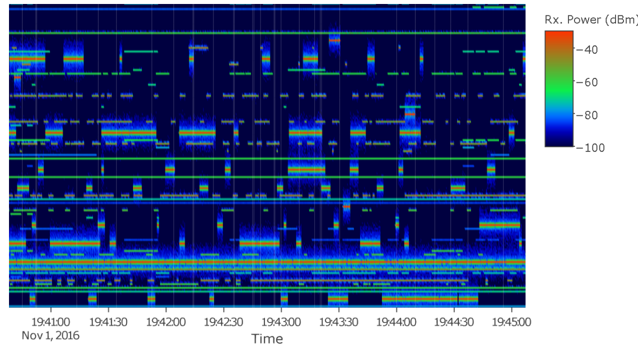


Fig. 3. A 5-minute snapshot of the received-power levels for LMR channels in 866-869 MHz.

to transmit control signalling required for the network's operation. The 0.3s time resolution of the measurements for each channel allows the temporal activity pattern of voice channels to be accurately characterized.

Fig. 4 shows a 10-hour snapshot (approximately 120,000 samples) of the temporal pattern of idle (burst-off) and busy (burst-on) states for a typical public safety voice channel. The empirical distributions of idle and busy state durations as well as log-normal fits for them (obtained via maximum likelihood estimation) are also shown in Fig. 4. We tried fitting a variety of common distributions such as exponential, log-normal, weibull, and pareto to the empirical data and found out that the log-normal distribution typically provided the best fit for trunked voice channels with many users, corroborating the results of [4].

Such data-driven modelling of incumbent users' spectrum activity patterns allows for more realistic simulation studies of spectrum sharing opportunities as we have shown in [2]. In the present work however, we focus on longer-term spectrum usage patterns which offer insights enabling a more proactive approach to spectrum management.

III. Spectrum Occupancy Analysis

We define spectrum occupancy as the percentage of time a frequency channel is detected to be in use with its received power being above the detection threshold as defined earlier. In this section we describe some of the occupancy patterns observed and how these might be leveraged for spectrum assignment. To facilitate the analysis over longer time periods, we generate aggregate *hourly* statistics from the high resolution measurement data. These include the hourly received-power histogram, idle and busy state duration histograms, and hourly occupancy of each channel, among others. Given the large size of high-resolution data (as shown in Fig. 3), continuous production of hourly statistics is performed on an Apache Spark cluster [5].

Figs. 5 and 6 show two typical plots of the hourly occupancy versus time for different LMR ranges for a

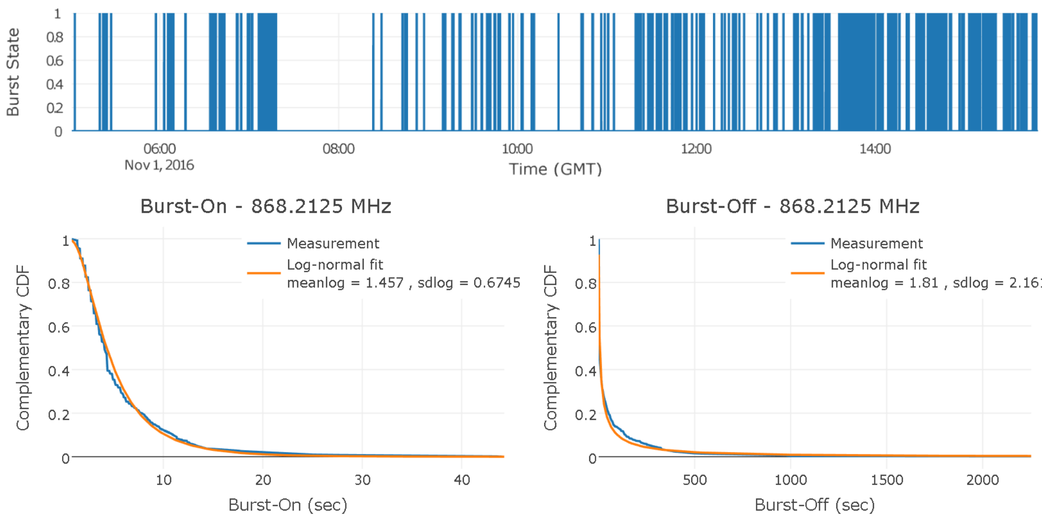


Fig. 4. Top: Channel idle (burst-off) and busy (burst-on) states pattern for a public safety voice channel over a 10-hour period; Bottom: Empirical distribution and log-normal fit for the idle and busy state durations

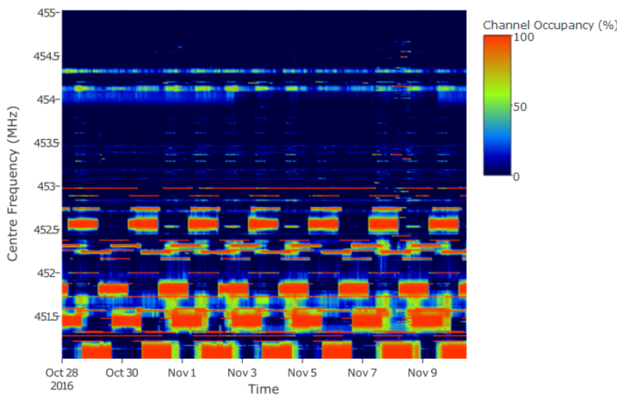


Fig. 5. Hourly occupancy of LMR channels in 451-455MHz range over a period of two weeks.

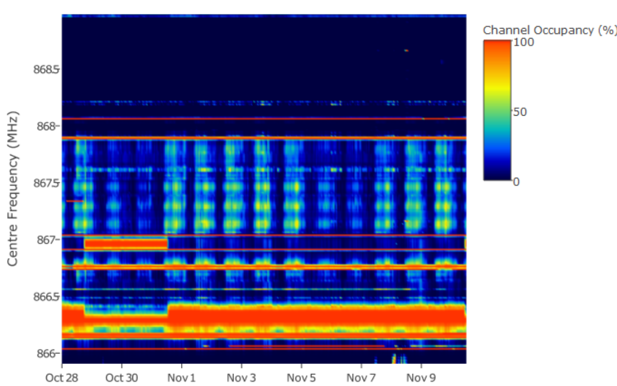


Fig. 6. Hourly occupancy of public safety LMR channels in 866-869MHz range over a period of two weeks.

2-week period. For display clarity, each plot has been limited to 4 MHz which still covers a large number of channels occupied by diverse LMR technologies and services, including commercial, public safety, government, and municipal public service users. The plots show that spectrum utilization is not uniform across the LMR bands, with some portions congested, while other portions are being either lightly used or completely unused. Some LMR traffic is observed to exhibit a periodic occupancy pattern that is repeated on a daily and weekly basis, with lower usage observed at night and on weekends. Such trends arise from social behaviour and habits and corroborate patterns observed for LMR use in prior work [6]. We also observe that in the Ottawa area, many LMR channels are not in use most of the time, both in the 450 MHz band and in the upper range of the 800 MHz band.

A. *Usage Pattern Seasonality and Prediction*

Inspection of the spectrum occupancy patterns in Fig. 6 suggests that hourly occupancy of LMR voice channels is both daily and weekly seasonal. Further analysis of the auto-correlation function (acf) for the hourly occupancy time-series, shown in Fig. 7, confirms this observation as the acf oscillates between its peaks and minima every 12 hours. Therefore, time series forecasting models such as seasonal ARIMA may be used to predict future occupancy patterns for these channels. For the aforementioned channels however, the augmented exponential smoothing of [7] was found to perform better. Unlike conventional seasonal ARIMA models, this model supports multiple seasonal periods (e.g., both daily and weekly patterns in our case) and allows the seasonality to vary slowly with time for a better fit.

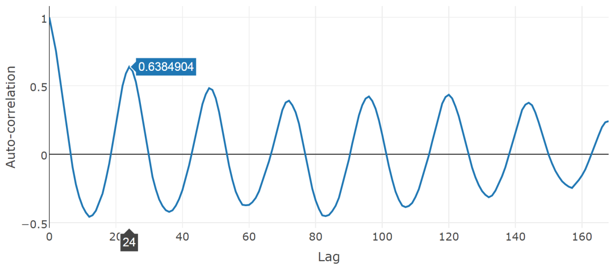


Fig. 7. Auto-correlation function of a typical LMR public safety voice channel.

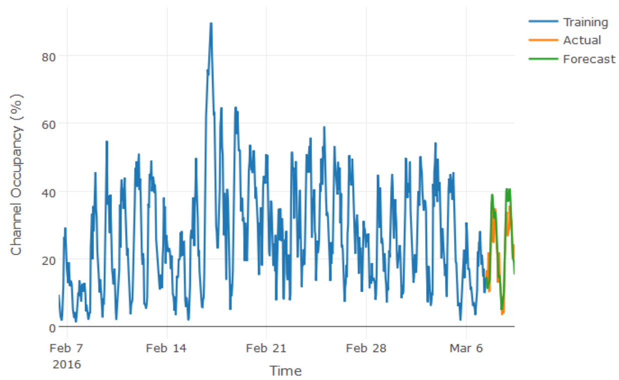
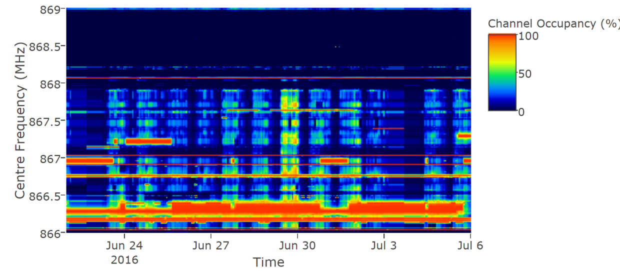
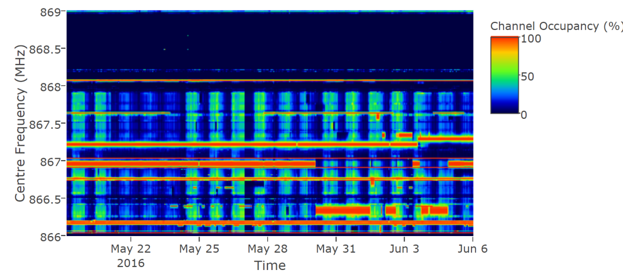


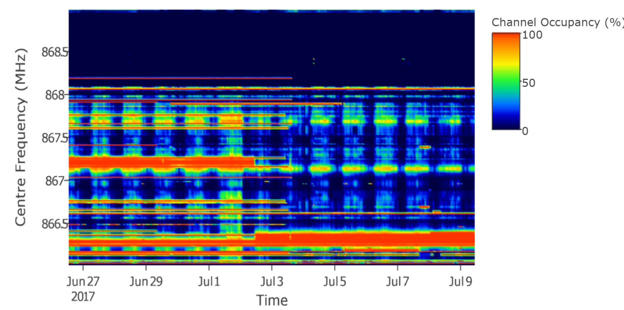
Fig. 8. 48-hour spectrum occupancy prediction for 867.8875MHz.



(a)



(b)



(c)

Fig. 9. Impact of special events on public safety LMR spectrum usage: (a) a visit by the US president (b) Ottawa Race Weekend (c) Canada Day 2017

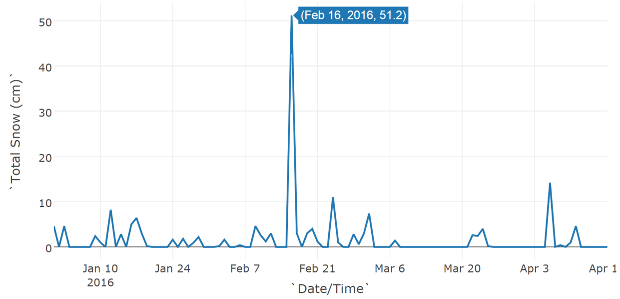
Fig. 8 shows an example of spectrum occupancy prediction using the above method for one of the public safety voice channels in Ottawa. In this case, one month of measurement data has been used as training to successfully predict the occupancy pattern over the next 48 hours. Having the ability to predict spectrum occupancy enables a more dynamic approach to spectrum assignment where spectral resources are assigned based on forecasted demand as opposed to over-provisioning for the worst-case scenario. For such predictive analytics to be viable however, it has to be able to account for occasional surges in the demand due to various external factors as we will discuss next.

B. Impact of External Factors

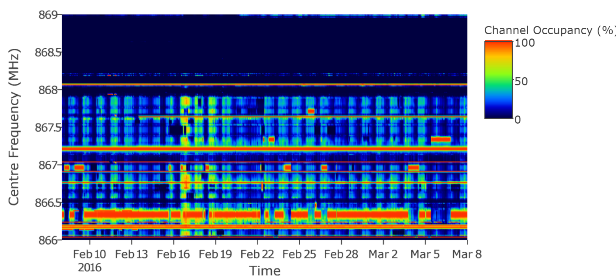
While daily and weekly patterns can be successfully used to model spectrum occupancy of most public safety LMR channels, there are occasional anomalies which can not be explained by the aforementioned seasonal effects. In many cases, such anomalies are due to external factors such as major events which drive up the communications needs of public safety users. Fig. 9 shows three examples of such anomalies: a) president of the United States visiting Ottawa on June 29, 2016, (b) Ottawa "Race Weekend" which includes many running races spread over two days (May 28-29th, 2016) and is the biggest event of its kind in Canada, (c) Canada Day on July 1st, 2017, which coincided with Canada's 150th birthday and drew very large crowds.

In all three cases, there were many road closures, changes to traffic patterns, and concentration of people in certain areas of the city core requiring extensive coordination and communications among public safety users. As illustrated in Fig. 9, there is a marked increase in spectrum occupancy compared to similar weekday or weekends for most public safety LMR channels. Fig. 9b for instance shows that the Race Weekend event on May 28-29th increased weekend spectrum usage almost to the level of a typical weekday. In case of Canada Day, with this being the largest Canada Day event in the history, there were extensive preparations in the days leading to Saturday, 1st of July, which is mirrored in increased occupancy for some LMR channels in the last week of June as seen in Fig. 9c.

Severe weather is another external factor affecting public safety spectrum demand. Fig. 10 shows how a snowstorm on February 16, 2016 (more than 50cm of snowfall in the span of 24 hours) triggered a few days of increased spectrum occupancy for Ottawa public



(a)



(b)

Fig. 10. Impact of severe weather on public safety LMR spectrum usage: (a) Total daily snowfall reported at Ottawa International airport (b) Hourly occupancy of Ottawa public safety LMR channels in 866-869MHz range

safety users. A similar observation was reported in [9], where the spectrum occupancy of police channels over regular days was compared with the occupancy during a blizzard. It is worth noting that features of weather data other than total daily snowfall shown in Fig. 10 may be more important. Specifically, we have observed instances of increased public safety LMR spectrum usage on days with only 3-5cm snowfall (which is fairly typical for Ottawa during winter). Further inspecting other weather attributes indicated blowing snow and low visibility which had resulted in numerous emergencies. Therefore, care must be taken when fusing spectrum measurements with other sources of data to include the most relevant features.

Authors in [8] also have studied the impact of major events on spectrum usage. In particular they modeled the *received-power* on a given LMR channel with a Gaussian Mixture Model (i.e. a mixture of Gaussian random variables a.k.a. GMM) and then showed that a model can be trained to distinguish event days by comparing the estimated parameters of the GMM model with those obtained on regular non-event days. This approach is predicated on there being changes to the received-power footprint due to big events. In the example presented in [8], the LMR channels belong to a transit company which runs additional buses and modified routes during major game events which can explain the variations in the received-power distributions. For many LMR channels though, the location of transmitters

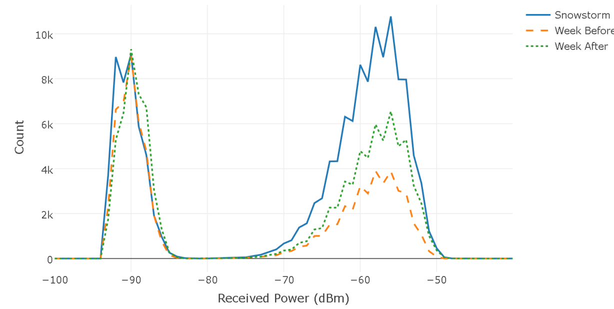


Fig. 11. Histogram of received power during a severe weather event (snowstorm) compared with those of regular days

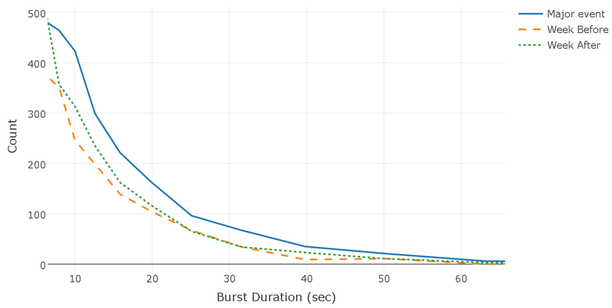
(and therefore shape of the received-power distribution) does not change from one day to another. This is illustrated in Fig. 11 which compares the measured daily histogram of received-power during a severe weather event (snowstorm) with those of regular days (7 days before and after). Given the two licensed transmitters on this channel are at fixed locations, their average received-power levels and shape of the distribution observed at the sensor do not change significantly. The intensity of the histogram however clearly indicates unusual activity for one of the transmitters due to severe weather. Therefore, the GMM approach used in [8] may need to be modified to use other *features* of the measurement data (e.g., occupancy as shown in Fig. 10).

Fig. 12 further shows an example of the impact of a major event or severe weather on the underlying characteristics of public safety users' activity. In particular, the right tail of the channel busy-duration histogram indicates a larger number of longer bursts which ultimately leads to a higher percentage of spectrum occupancy (as observed in Figs. 9 and 10b). Understanding the impact of external factors on the distribution of channel busy and idle durations allows more accurate modeling of spectrum sharing scenarios, where decisions on sharing a channel would depend on the characteristics of the gaps in spectrum activity [2].

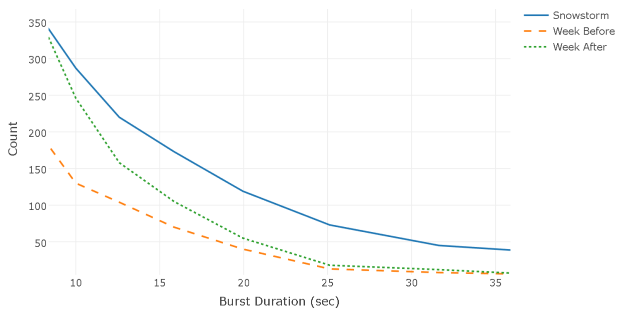
IV. Conclusion

In this work, we presented preliminary exploratory analysis of LMR spectrum measurement data collected in Ottawa, Canada, which was shown to have distinct time and frequency patterns which could be used to train predictive models. We further illustrated existence of correlations between spectrum occupancy of public safety users and external factors, such as major events and weather.

The ability to predict future spectrum occupancy while taking into account other sources of data, such as weather forecast and upcoming events, allows a more proactive approach to spectrum assignment. Under such a paradigm, capacity bottlenecks are predicted and managed, before they occur, with minimum human



(a)



(b)

Fig. 12. Tail of the histogram for burst (channel busy) duration compared on regular days with that of (a) a major event (U.S. president visit) (b) a severe weather event (snowstorm)

intervention. As the size of spectrum measurement data grows, so does our understanding of the fundamental factors driving the demand for various users and services, thereby enabling us to transition toward a paradigm where spectrum shortage and oversupply across different networks can be predicted and managed proactively.

Results presented herein were based on data obtained from a single sensor. Given the availability of spatial data from the network of sensors shown in Fig. 1, future work will explore the spatial properties of the measurements with potential applications to data-driven propagation models, interference management, and spectrum sharing.

Acknowledgment

The author would like to thank Kareem Baddour, Paul Guinand, and Humphrey Rutagemwa for helpful discussions and feedback during the course of this work.

References

[1] L. Li *et al.*, ”A cloud-based spectrum environment awareness system,” *in proc. IEEE International Symposium on Personal, Indoor, Mobile Radio Communications (PIMRC)*, Montreal, Canada, October 2017.

[2] H. Rutagemwa, K. E. Baddour, A. Ghasemi, ”Spectrum sharing opportunities in land mobile radio bands: A data-driven approach,” *in proc. IEEE PIMRC Workshop on Cognitive Radio and Innovative Spectrum Sharing Paradigms for Future Networks*, Montreal, Canada, October 2017.

[3] S. Wang, F. Patenaude, R. Inkol, ”Upper and lower bounds for the threshold of the FFT filter bank-based summation CFAR detector,” *in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2006)*, pp. 289-292, May 2006.

[4] T. Taher, R. Bacchus, K. Zdunek, D. Roberson, ”Empirical modeling of public safety voice traffic in the land mobile radio band,” *in. Proc. Int. Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, 2012.

[5] Apache Spark: Lightning-fast cluster computing, https://spark.apache.org/

[6] T. Taher, R. Bacchus, K. Zdunek, D. Roberson, ”Long-term spectral occupancy findings in Chicago,” *in Proc. IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DYSPAN)*, pp. 100- 107, June 2011.

[7] A. M. De Livera and R. J. Hyndman, ”Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing,” *Journal of the American Statistical Association*, vol. 106, no. 496, pp. 1513-1527, 2011.

[8] A. Abdallah *et al.*, ”Detecting the Impact of Human Mega-Events on Spectrum Usage,” *13th IEEE Annual Consumer Communications and Networking Conference (CCNC)*, Las Vegas, January 2016.

[9] T. Taher, R. Bacchus, K. Zdunek, D. A. Roberson, ”Dynamic spectrum access opportunities for public safety in land mobile radio bands,” *in proc. Int. Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, June 2011.