Semantic Models for Labeling Spectrum Data

Vaishali Nagpure(✉), Cynthia Hood, and Stephanie Vaccaro

Illinois Institute of Technology, Chicago, IL 60616, USA
{vnagpure,svaccaro}@hawk.iit.edu, hood@iit.edu

Abstract. With the increasing importance and demand for wireless communications, there is a spectrum crunch and there is no easy way to meet this increasing demand. Spectrum monitoring is a key enabler for understanding how the spectrum is currently used and identifying opportunities for spectrum sharing. This paper focuses on analysis of a large volume of complex spectrum measurements in the public safety band and involves modeling spectrum sensors, public safety domain knowledge and events involving public safety. The information models developed can be used to create labels for the complex spectrum data and facilitate the analysis of the data needed for spectrum sharing.

1 Introduction

Wireless communication plays an increasingly important role across a variety of domains. With the growing demand for new and expanded wireless services, one of the key challenges is the scarcity of spectrum. Spectrum is a finite resource and in the United States (as in most parts of the world) the spectrum has been fully allocated. This means that there is no "free" spectrum that can be utilized to meet the growing demands, thereby making spectrum sharing necessary.

An important step to determine the feasibility of sharing is to learn how the spectrum is currently being used. This involves understanding the occupancy of the spectrum in space, time and frequency and identifying where opportunities exist for sharing. Characterizing the details of how the spectrum is used – e.g., the time-frequency utilization patterns is fundamental to the design of dynamic spectrum access (DSA) systems as it can significantly increase the opportunity for spectrum sharing by leveraging signal periodicity. Obtaining detailed spectrum utilization patterns is challenging, particularly when analyzing multiple bands.

Multiband spectrum measurements are highly complex spatio-temporal data sets that require very specialized domain knowledge to collect, analyze and interpret. [1] The entire process is extremely time consuming. To analyze and interpret the measurements, a variety of contextual information ranging from the specific configuration of the spectrum sensor and potential emitters to drivers of spectrum use is needed. Current analysis is largely manual, and typically limited to one or a few bands. Multiband analysis is usually based on visual interpretation of waterfall charts or spectrograms. Although context is sometimes used anecdotally to interpret the dynamics of spectrum usage, it has not been integrated into the characterization. Deciding what is normal or abnormal

L. Iliadis et al. (Eds.): AIAI 2018, IFIP AICT 520, pp. 3–12, 2018.
https://doi.org/10.1007/978-3-319-92016-0_1

for each band requires domain expertise. These factors limit the amount of analysis that is done on spectrum measurements and thus limit the understanding of existing spectrum use.

This work focuses on the analysis of spectrum measurements in the Land Mobile Radio (LMR) bands that include the spectrum used for public safety in Chicago. Given the nature of public safety communication, it is important to be able to meet communication demands whenever emergencies or other large-scale events occur. Therefore it is critical to understand how the public safety spectrum is used in dynamic situations, especially as these events are unfolding. More specifically, how do the communication demands evolve over time and what triggers the demands. To understand the triggers, it is necessary to understand the context of the demands.

Figure 1 was generated from two weeks of observation of Land Mobile Radio (LMR) band data in 2009 and shows the difference in spectrum occupancy on weekdays and weekends. Figure 2 shows the increase in call rate in the 406 MHz LMR band during the NATO Summit held in Chicago in May 2012. Both of these figures show that there is a relationship between spectrum use and human activity.

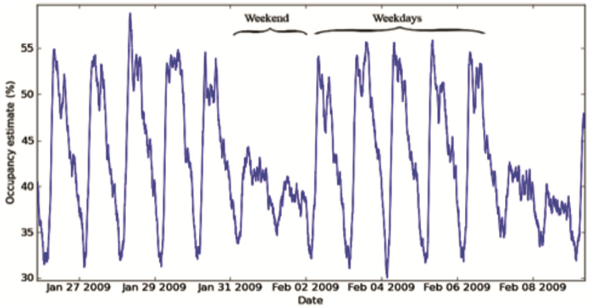


Fig. 1. Two week plot of spectrum occupancy in the 450 MHz LMR Band

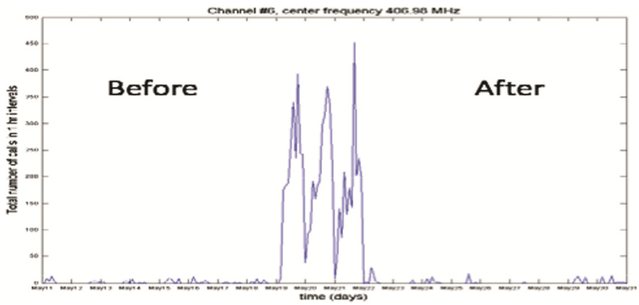


Fig. 2. Call rate in 406 MHz band

As illustrated above, context is necessary to fully understand the significance of the patterns and anomalies in the spectrum measurements. More importantly, these patterns and events need to be identified so they can be incorporated into demand models. In an urban environment like Chicago, events drive a significant portion of public safety communication. These events may be planned (i.e. sporting or entertainment events) or unplanned emergency events of different scales and at different locations.

The importance and role of human activity has been recognized in context-aware services and networks. Context enhances understanding of spectral usage through improvement in the process of spectrum data analysis as well as interpretation of the results. One of the challenges with wideband monitoring efforts is identifying events or time intervals of interest to analyze. Context may be utilized to identify potential events of interest for near-term or offline analysis. As a first step toward automation of spectrum analysis, this paper proposes to model configuration, domain knowledge and other potentially relevant information with the use of semantic modeling in a way that it can be fused with measurements for analysis and in particular can provide labels for the spectrum data.

The rest of the paper is organized as follows. Section 2 describes the challenges of measuring the spectrum and provides details on the measurements utilized in this work and typical analysis techniques. Public safety communication and related work is described in Sect. 3. Our approach and a case study during the Chicago Marathon is described in Sect. 4. Finally, conclusions and future work is discussed in Sect. 5.

2 Spectrum Measurements

The spectrum is organized into bands where a band is a contiguous range of frequencies that are used for the same general purpose. Bands may be further subdivided into channels. The frequency resolution is the sampling interval in the frequency domain whereas the resolution bandwidth (RBW) is the bandwidth of a single spectrum measurement obtained during a specific sampling interval in time at a specific location in the spatial domain. If the RBW is too large, part of the adjacent channel power is measured during the measurement of a certain channel leading to incorrect results. This is an enormously challenging constraint and may lead to the use of multiple detectors, each covering a specific portion of the spectrum to provide a small enough RBW to ensure that all channels are correctly registered and at the same time the revisit time is positioned at a level where no signals are missed.

The band plan is also a critical part that is basically a division of the particular range of radio frequencies of interest into bands with common parameter settings such as resolution bandwidth. Typically band plans are relatively static, but if you need to capture details with higher resolution measurements on a particular day for an event, the band plan can be dynamic instead of static. For example, if we are aware of an event such as a sporting event or protest, the band plan can be changed to gather higher resolution measurements in bands that might be impacted (e.g. public safety).

The IIT Spectrum Observatory has been monitoring the 30–6000 MHz radio spectrum of the city of Chicago since mid-2007 from its location on top of the 22 story Tower on IIT's Main campus on the south side of Chicago. [2] Energy detection sensing is used in our measurement system and the resolution bandwidth is kept at 3 kHz for all bands. This results in around 93 MB of data per day.

In the case of the public safety bands where channel bandwidths are narrow (<30 kHz), due to the high sweep time, short transmissions cannot be detected. To address these issues, a new measurement system utilizing a Tektronix RSA 306 spectrum analyzer was added. This allows us to obtain higher time and frequency resolution data with a focus on the LMR bands which include the public safety bands in Chicago. After conducting experiments, resolution bandwidth of 1 kHz for the Chicago Police Department (CPD) frequencies is configured as CPD is a case study for this research. This results in around 16 GB of data per day. The measurements are transmitted to a local storage system daily.

2.1 Analysis with Spectrograms

The data files produced are transmitted from the data capture PC to a high performance workstation. The data is processed to produce a variety of plots and usage statistics. Spectrograms or waterfall charts are used to present much of the spectrum occupancy data. For these charts, time is represented on the y-axis (days or hours), frequency on the x-axis (MHz) and power (dBm) as color level, with the indicated mapping. Higher power indicates higher occupancy. Several waterfall charts derived shown below in Figs. 3, 4 and 5.

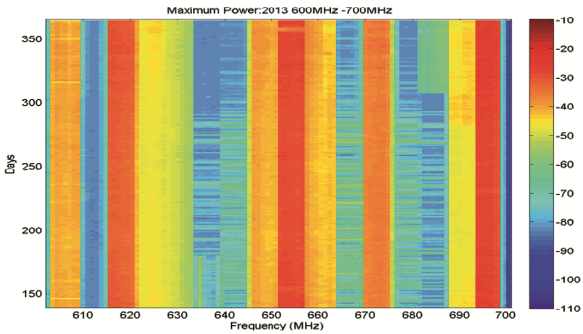


Fig. 3. 600 MHz television bands during 2013

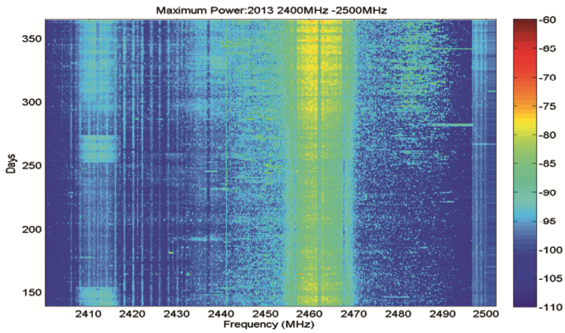


Fig. 4. 2.4 GHz Band during 2013

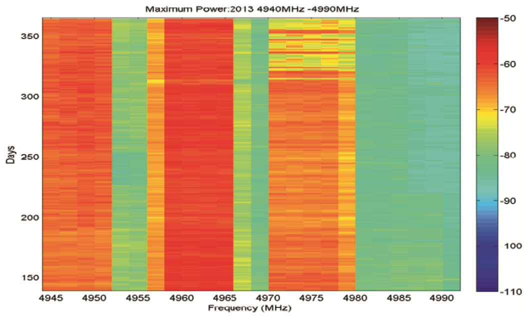


Fig. 5. 4.945–4.99 GHz Broadband Public Safety during 2013

The spectral occupancy was estimated based on the threshold method and different thresholds were selected for each band by estimating the noise floor of the system. The utilization of different bands can be found to range from highly utilized through sporadically used to not used at all. A good example of highly utilized bands is the TV bands shown in Fig. 3 with their consistently high power. The obvious reason for the high occupancy is that TV channels in a metropolitan market like Chicago are always broadcasting and are well occupied. The interpretation of Fig. 3 is quite straightforward once you know how the band has been allocated. Notice that Figs. 4 and 5 look different and explaining the visualizations shown in these figures is more challenging. It is clear that one must be familiar with the band allocations, the measurement system location and configuration, as well as usage behavior within the band to accurately interpret the measurements.

3 Public Safety Background and Related Work

Public safety communication networks are typically designed with enough capacity to handle "worst case" scenarios, with some channels largely unused under typical day-to-day activity [3]. Studies [4] have noted this to be the case with occupancy figures generally being fairly low, however during atypical periods of high activity – most notably during disasters and other emergencies – channel capacities may approach or even meet their limits.

There have been many studies of public safety traffic such as [5–8] and including [9] utilizing the SDR-based spectrum monitoring system mentioned above to collect measurements. These studies are focused on traffic in the deployed LMR systems. They generally consider typical traffic along with traffic during large-scale disaster or emergency events and generally separate these into two classes. Part of the challenge with having richer models is the lack of events to study [10].

Patterns of communication are identified [6] and models of heterogeneous networks are developed [11] and mobility on disaster sites is considered. [12] Accurate traffic models are necessary for a service provider of emergency communications to properly maintain the capacity planning of the network. To do so, good traffic models need to be developed that can capture characteristics including occurrences of few and large random incidents and accidents. These occurrences can be described as unusual spikes and/or long tails in a probability model term of an actual network load.

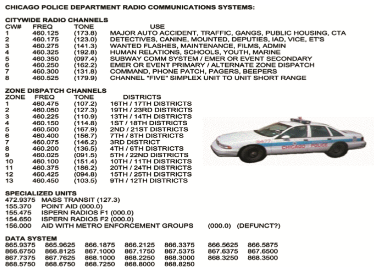
Today's Public Safety agencies and organizations have started planning to evolve their networks to LTE based public safety solutions [13]. With LTE, the first responders can access a wide variety of services, starting from high bandwidth to real time communications. This will help the first responders in case of mission critical communications such as natural disasters or terrorist attacks. In [14], a virtual Public Safety (PS) operator is proposed that relies on shared infrastructure of commercial LTE networks to deliver services to its users. Several methods of allocating spectrum resources between virtual operators are compared at peak times and they examine how this influences differing traffic services. They showed that it is possible to provide services to the Public Safety users reliably during both normal and emergency operation. The authors of [12] focused on the performance evaluation of the communication systems in disaster scenarios.

4 Approach and Case Study

One of the challenges with wideband spectrum measurements, particularly long-term measurements, is finding data of interest. For example, the IIT Spectrum Observatory collects wideband measurements between 30–6000 MHz. This covers many different frequency bands and channels over a wide variety of uses. There is a wealth of data over many years, so how to choose a frequency band and time period to study? In the context of dynamic spectrum sharing, the focus is on frequency bands that are not fully utilized and we are generally interested in understanding the usage and how the usage changes over various time periods. But it is not enough to understand changes in the energy from a signal processing perspective, it is also important to understand why the usage is changing. This is relatively straightforward for major events like the NATO Summit shown in Fig. 2, because you can identify the days of interest and then visualize the data. This can be quite challenging for many other types of events though given the manual nature of visualization. Patterns can also be identified manually through visualization, but for the same reasons can be challenging to capture without time-consuming visualization. This can be effective for small datasets but does not scale well. To address this issue, this paper proposes to combine different types of information from several different sources using both statistical and semantic modeling techniques.

4.1 Case Study

A large city like Chicago has many different types of events occurring almost on a daily basis providing a good opportunity to model spectrum use over a variety of different scenarios. This case study focuses on the channels used by the Chicago Police Department (CPD). The goal of this case study is to step through the analysis of an event that results in noticeable changes to identify the information and steps needed for automation. The event we studied was the Chicago Marathon on October 8, 2017. The Chicago Marathon has a 26.2 mile course that covers 22 neighborhoods in Chicago. Approximately 45,000 runners participate in the marathon and there are many fans along the way. Public safety works to secure the course as well as to do traffic management around the city. In this case study, as in the past, we traversed backwards from events of interest to see if the impact on spectrum could be detected. The first question was what frequencies to analyze. The frequencies used by CPD are shown in Fig. 7. CPD has organized the city into zones that cover one or more districts as shown in Fig. 6. [15] The Chicago neighborhood and district map is shown in Fig. 7. [15].



CHICAGO POLICE DEPARTMENT RADIO COMMUNICATIONS SYSTEMS:

CITYWIDE RADIO CHANNELS

CW#	FREQ	TONE	USE
1	460.125	(173.8)	MAJOR AUTO ACCIDENT, TRAFFIC, GANGS, PUBLIC HOUSING, CTA
2	460.175	(123.0)	DETECTIVES, CANINE, MOUNTED, DEPUTIES, IAD, VICE, ET'S
3	460.275	(141.3)	WANTED FLASHES, MAINTENANCE, FILMS, ADMIN
4	460.325	(192.8)	HUMAN RELATIONS, SCHOOLS, YOUTH, MARINE
5	460.350	(097.4)	SUBWAY COMM SYSTEM / EMER OR EVENT SECONDARY
6	460.250	(162.2)	EMER OR EVENT PRIMARY / ALTERNATE ZONE DISPATCH
7	460.300	(131.8)	COMMAND, PHONE PATCH, PAGERS, BEEPERS
8	460.525	(179.9)	CHANNEL "FIVE" SIMPLEX UNIT TO UNIT SHORT RANGE

ZONE DISPATCH CHANNELS

ZONE	FREQ	TONE	DISTRICTS
1	460.475	(107.2)	16TH / 17TH DISTRICTS
2	460.050	(127.3)	19TH / 23RD DISTRICTS
3	460.225	(110.9)	13TH / 14TH DISTRICTS
4	460.150	(114.8)	1ST / 18TH DISTRICTS
5	460.500	(167.9)	2ND / 21ST DISTRICTS
6	460.400	(156.7)	7TH / 8TH DISTRICTS
7	460.075	(146.2)	3RD DISTRICT
8	460.200	(136.5)	4TH / 6TH DISTRICTS
9	460.025	(091.5)	5TH / 22ND DISTRICTS
10	460.100	(151.4)	10TH / 11TH DISTRICTS
11	460.375	(186.2)	20TH / 24TH DISTRICTS
12	460.425	(094.8)	15TH / 25TH DISTRICTS
13	460.450	(103.5)	9TH / 12TH DISTRICTS

SPECIALIZED UNITS

472.9375 MASS TRANSIT (127.3)
155.370 POINT AID (000.0)
155.475 ISPERN RADIOS F1 (000.0)
154.650 ISPERN RADIOS F2 (000.0)
156.000 AID WITH METRO ENFORCEMENT GROUPS (000.0) (DEFUNCT?)

DATA SYSTEM

865.9375	865.9625	866.1875	866.2125	866.3375	866.5625	866.5875
866.6750	866.8125	867.1000	867.1750	867.5375	867.6375	867.6500
867.7375	867.7625	868.1000	868.2250	868.3000	868.3250	868.3500
868.5750	868.6750	868.7250	868.8000	868.8250		

Fig. 6. Frequencies used by the Chicago Police Department [15]

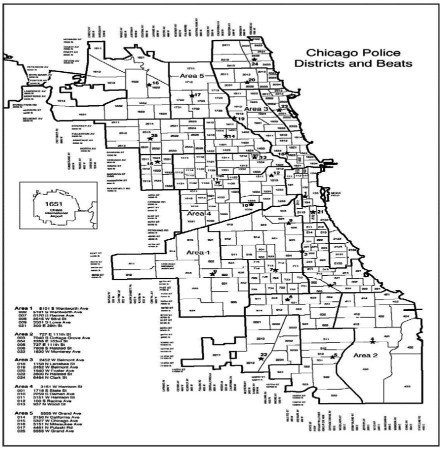


Fig. 7. Chicago Police Department Radio Zone Map [15]

As shown in Fig. 6, there are both citywide channels for specific types of communication as well as location-specific zone channels. To determine the communication channel and frequency for a given public safety event, it is necessary to determine the type of event and also the location. Each location must be translated into the appropriate district and corresponding zone to determine the channel and corresponding frequency of communication related to the event.

The relationship between neighborhoods, districts and zones shown in Fig. 7 along with the CPD radio communication information was encoded using SWI-Prolog. Events from various City of Chicago event calendars were also encoded. For each event, the neighborhood was included. The result of a query to the resulting knowledge base for the date “October 8, 2017” is shown in Fig. 8. The marathon goes through several neighborhoods so the corresponding zones are listed. In addition, since the marathon is considered an “event” and citywide channels 1,5 and 6 are used for events, they are also listed.

Fig. 8. Result of the query to knowledge database.

We analyzed the bands corresponding to the zone and citywide channels in the query result. Citywide channel 6 showed unusual occupancy. The waterfall charts for citywide channel 6 are shown in Figs. 9 and 10. Figure 9 shows a sweep time from 11 pm on October 7, 2017 to 11 pm on October 8, 2017. Figure 10 shows from 11 pm on October 8, 2017 to 11 pm on October 9, 2017. These waterfall charts plot the energy detected in the band on each sweep. The x-axis gives the frequency and the y-axis gives the sweep number. There is one sweep approximately every 2 s. The channel beginning at 460.25 MHz is clearly evident from the waterfall chart. The yellow color indicates a higher level of power in the band. It is clear from Fig. 9 that there is activity in the band starting very early in the morning of October 8th and continuing through 11 pm. Looking at Fig. 10, one can see the activity continuing until the early hours of October 9th. This appears to be consistent with expected CPD activity during the marathon.

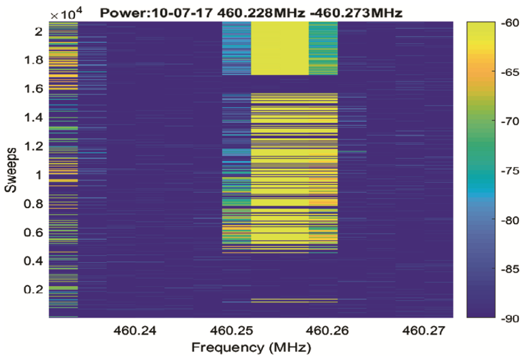


Fig. 9. Waterfall chart for Chicago Police Citywide channel 6 (460.25 MHz)

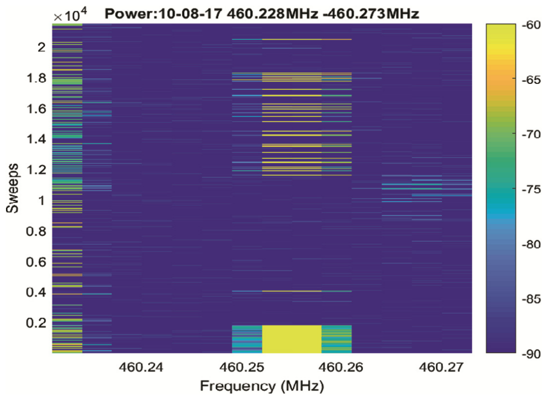


Fig. 10. Waterfall chart for Chicago Police Citywide channel 6 (460.25 MHz)

It is clear that a significant amount of information is needed to put together this simple analysis. One of the reasons we chose to study the public safety spectrum is that there are many publicly available sources of information that are relevant. This information and the relationships between various pieces of information can be modeled using semantics. These models can be used to reason about potential explanations for increased activity given unusual spectral activity. It can also be used to trigger analysis of specific frequency bands due to a certain type of event.

5 Conclusion and Future Work

This paper describes the challenges of spectrum measurement analysis and motivates the need for more in-depth analysis and fine-grained modeling for the purpose of spectrum sharing. The necessary analysis is limited by the complexity of the measurement, analysis and interpretation of the data. Analysis requires a substantial amount of domain knowledge along with contextual information. Contextual information can be derived from a variety of different information sources including events.

This paper focuses on the public safety spectrum in the LMR band in Chicago. More specifically, we study the CPD spectrum. Semantic models were developed and coded in SWI-Prolog to describe the organization and use of the CPD allocated spectrum along with planned events in Chicago. The models were used to help identify relevant channels used during the Chicago marathon. These models can be used to label periods of unusual activity. Our case study uses visualization of the waterfall charts to determine unusual activity. The contribution of this work is the use of semantic models for labeling. Our long-term goal is to automate the analysis of spectrum data across frequency and time.

Ongoing work aims to create statistical models of typical activity and identify periods of unusual activity automatically. These periods of unusual activity can then be correlated with events or vice versa. Public Safety spectrum behavior is challenging to interpret and prediction of usage is driven by many factors such as planned and unplanned events, weather and human protocols. Further development of the semantic models is needed to capture the many different information sources as well as the correlations across time and frequency bands.

References

1. Ding, G., Wu, Q. Wang, J., Yao, Y.-D.: Big Spectrum Data: The New Resource for Cognitive Wireless Networking, April 2014. http://arxiv.org/pdf/1404.6508.pdf
2. Bacchus, R.B., Fertner, A.J., Hood, C.S., Roberson, D.A.: Long-term, wide-band spectral monitoring in support of dynamic spectrum access networks at the IIT spectrum observatory. In: 3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2008, pp. 1–10. IEEE (2008)
3. Lehr, W., Jesuale, N.: Public safety radios need to pool spectrum. IEEE Commun. Mag. **47**(3), 103–109 (2009)
4. Bacchus, R., Taher, T., Zdunek, K., Roberson, D.: Spectrum utilization study in support of dynamic spectrum access for public safety. In: 2010 IEEE Symposium on New Frontiers in Dynamic Spectrum, pp. 1–11. IEEE (2010)
5. Aschenbruck, N., Gerharz, M., Frank, M., Martini, P.: Modelling voice communication in disaster area scenarios. In: Proceedings of IEEE Conference on Local Computer Networks (LCN2006), pp. 211–220 (2006)
6. Aschenbruck, N., Gerharz, M., Martini, P.: How to assign traffic sources to nodes in disaster area scenarios. In: Proceedings of the 26th IEEE International Performance Computing and Communications Conference (IPCCC) (2007)
7. Sharp, D.S., Cackov, N., Lasković, N., Shao, Q., Trajković, L.: Analysis of public safety traffic on trunked land mobile radio systems. IEEE J.-SAC **22**(7), 1197–1205 (2004)
8. Vujicić, B., Cackov, N., Vujicić, S., Trajković, L.: Modeling and characterization of traffic in public safety wireless networks. In: Proceedings of SPECTS, pp. 214–223 (2005)
9. Taher, T.M., Bacchus, R.B., Zdunek, K.J., Roberson, D.A.: Empirical modeling of public safety voice traffic in the land mobile radio band. In: 2012 7th International ICST Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), pp. 230–235. IEEE (2012)
10. Chen, C., Pomalaza-Raez, C., Colone, M., Martin, R., Isaacs, J.: Modeling of a public safety communication system for emergency response. In: 2009 IEEE International Conference on Communications, Dresden, pp. 1–5 (2009)
11. Badarch, T., Bataa, O.: Incoming traffic modeling of heterogeneous public safety network. In: 2013 15th Asia-Pacific Network Operations and Management Symposium (APNOMS), Hiroshima, Japan, pp. 1–6 (2013)
12. Aschenbruck, N., Martini, P.: Modeling public safety scenarios to evaluate wireless communication systems. In: 2009 1st International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology, Aalborg, pp. 510–514 (2009)
13. LTE networks for Public Safety Services, Nokia Networks, White Paper
14. van de Belt, J., Ahmadi, H., Doyle, L.E., Sallent, O.: A prioritized traffic embedding mechanism enabling a public safety virtual operator. In: Proceedings of IEEE 82nd Vehicular Technology Conference (VTC2015-Fall) (2015)
15. https://www.radioreference.com. Accessed 15 Mar 2018