Demo: Recommendation System for Dynamic Spectrum Access Through Spectrum Mining and Prediction

Ying Dai
Department of Computer and Information Sciences
Temple University
Philadelphia, PA 19122
ying.dai@temple.edu

Andrew Daniels
Department of Computer and Information Sciences
Temple University
Philadelphia, PA 19122
andrew.s.daniels@temple.edu

Jie Wu
Department of Computer and Information Sciences
Temple University
Philadelphia, PA 19122
jiewu@temple.edu

ABSTRACT

Dynamic spectrum access (DSA) [2] is a promising solution for improving the efficiency of channel utilization. The current spectrum allocation is static with many spectrum resources that are used very infrequently and are wasted. The DSA enables secondary users, or nodes, to transmit data packets on channels that are assigned to primary users (PUs). The constraint here is to make sure that PUs are not interfered with by SUs. Fig. 1 is an example with nodes that can make use of the spectrum when it is not occupied by the PU.

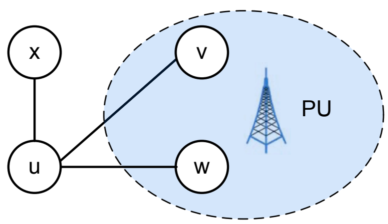

Figure 1: An example of DSA.

To make sure the PUs are not interfered with, one key component is the spectrum sensing. Many works have been done on spectrum sensing [5]. However, among the overall spectrums, white spaces are the parts whose usage can be exploited at specific times and locations. These patterns reduce the necessity of spectrum sensing, and motivate the database-driven schemes [4]. The IEEE 802.11af [3] defines the standard for spectrum sharing on white spaces. It relies on the database to provide white space maps over time at given locations. Every node can use the spectrum based on the given white space map. However, it is difficult and impractical to provide or construct such a spectrum database for all spectrums.

Inspired by the spectrum database in the white spaces, we propose a recommendation whose function is similar to that of the spectrum database. We aim at predicting the spectrum availabilities and recommending the part that is more likely to be available for the usage of SUs. The prediction on the spectrum availabilities is through running the spectrum mining algorithm on the collected spectrum data. Therefore, the two main components for our demo are the data collection and spectrum mining, as shown in Fig. 2. After the mining process is finished, our system would give the predictions and recommendations on the spectrum that is more likely to be available.

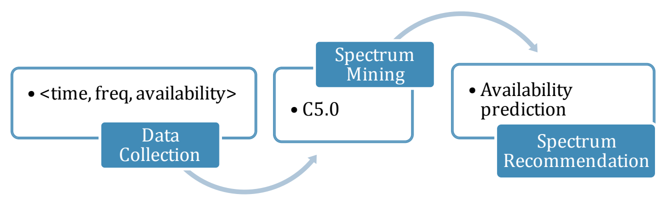

Figure 2: The process of our recommendation system.

The data collection is performed by the Universal Software Radio Peripheral (USRP) platforms, with two daughterboards (RFX 1200 and Basic RX). Due to the reason that the availability is both time- and frequency-sensitive, each data entry within our collected results contains three dimensional information: time, spectrum, and availability. The time here is the system time. The spectrum is the central frequency of each channel we divided. The availability is determined by setting a threshold on the SINR value on the receiver's side.

Currently, for the RFX 1200 daughterboard, we have 7 channels and each has a 40 MHz bandwidth. For the Basic RX daughterboard, we have 2 channels and each has a 100 MHz bandwidth. The overview of the testbed for data collection is shown in Fig. 3. We use the corresponding senders (PUs) with RFX 1200 and Basic TX on the other side to generate the data, due to the limitation of the antenna ranges. Our program automatically switches among the subchannels. For every subchannel, it stays for three seconds to sense the signal strength.

We make use of the tools for See5 and C5.0[1] to mine the rules in the collected data set. The original collected data entry in the file contains the <time, central frequency, availability>. The collected data sample is shown in Fig. 4. The time is the system. Every 3 seconds, the receiver automatically jumps to the next channel. The availability is

MobiHoc'14, August 11–14, 2014, Philadelphia, PA, USA.
ACM 978-1-4503-2620-9/14/08.
http://dx.doi.org/10.1145/2632951.2636059.

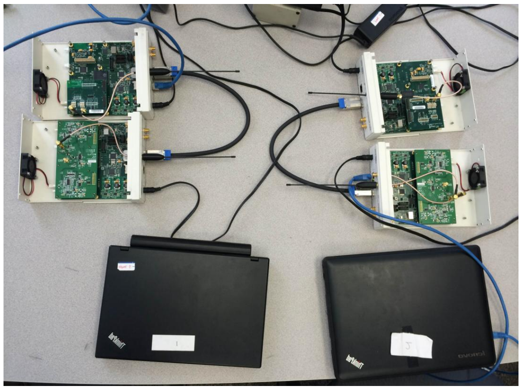

Figure 3: The overview of the data collection.

a binary value, by setting a threshold to the SINR values. The SINR values are calculated as the average value during the 3 seconds. The availability is set as the targeted value, which needs to be predicted for the future. Currently, our dataset is collected in 1 hour, which will be extended in the future. The generated rules are used to predict the channel availabilities given a specific time and frequency information. The location here is static in our testbed due to the antenna range limitation. We will extend it to different locations and add the location dimension to the dataset.

```
2014-05-30 16:18:25,1280000000,1
2014-05-30 16:18:28,1300000000,1
2014-05-30 16:18:31,1320000000,1
2014-05-30 16:18:34,1340000000,0
2014-05-30 16:18:37,1360000000,1
2014-05-30 16:18:40,1380000000,0
2014-05-30 16:18:43,1400000000,0
2014-05-30 16:18:46,1420000000,0
2014-05-30 16:18:49,1440000000,0
2014-05-30 16:18:52,1160000000,1
2014-05-30 16:18:55,1180000000,0
```

Figure 4: The screenshot of the collected data sample.

The other specifications for our demo are:

- Equipments to be used for the demo: Four USRP N200 with four daughterboards (RFX1200, Basic TX/RX); Two laptops.
- Space needed and the required setup time: $2 \times 2\ m^2$; 20-30 mins.
- Additional facilities needed: Six power supplies.

Through our demo, we show the potential of building the spectrum recommendation system for any spectrum range. Instead of the current sensing-only or database-driven schemes, the recommendation system can be a hybrid approach of the sensing-only and database-driven schemes. Every node can receive the recommendation on the spectrum that is more likely to be available, and can sense the spectrum to make sure the PUs are inactive. In this way, the sensing efficiency can be improved, and the PUs can be better protected. The mining part can be deployed in a distributed way, either on the selected nodes or new components that are similar to the current location-based databases.

Our future work includes adding the location dimension to our dataset. The limited space and antenna range makes it impractical to include location information in the current data collection. Also, we are going to explore the performance differences among different parameter settings, including the sensing time, channel switching time, and the number of subchannels.

1. REFERENCES

[1] Data mining tools see5 and c5.0, http://www.rulequest.com/see5-info.html.

[2] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty. Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey. *Computer Networks*, 2006.

[3] A. Flores, R. Guerra, E. Knightly, P. Ecclesine, and S. Pandey. Ieee 802.11af: a standard for tv white space spectrum sharing. *IEEE Communications Magazine*, 2013.

[4] R. Murty, R. Chandra, T. Moscibroda, and P. Bahl. Senseless: A database-driven white spaces network. *IEEE Transactions on Mobile Computing*, 2012.

[5] T. Yucek and H. Arslan. A survey of spectrum sensing algorithms for cognitive radio applications. *IEEE Communications Surveys Tutorials*, 2009.