Crowdsourced Misuse Detection in Dynamic Spectrum Sharing Wireless Networks

Debarun Das
School of Computing and Information
University of Pittsburgh
Pittsburgh, USA
Email: ded59@pitt.edu

Taieb Znati
School of Computing and Information
University of Pittsburgh
Pittsburgh, USA
Email: znati@pitt.edu

Martin Weiss
School of Computing and Information
University of Pittsburgh
Pittsburgh, USA
Email: mbw@pitt.edu

Pedro Bustamante
School of Computing and Information
University of Pittsburgh
Pittsburgh, USA
Email: pjb63@pitt.edu

Marcela M. Gomez
School of Computing and Information
University of Pittsburgh
Pittsburgh, USA
Email: mmg62@pitt.edu

J. Stephanie Rose
School of Computing and Information
University of Pittsburgh
Pittsburgh, USA
Email: jsr67@pitt.edu

***Abstract*— This paper proposes a spectrum enforcement framework by mobile, crowdsourced agents, who work in collaboration with a trustworthy infrastructure. To address the scarcity of spectrum, the Federal Communications Commission (FCC) mandated dynamic sharing of spectrum among the different tiers of users. The success of spectrum sharing, however, relies on the automated enforcement of spectrum policies. While most works in the past focus on automating spectrum enforcement before an actual harm has occurred, we focus on *ex post* spectrum enforcement, which happens during/after the occurrence of a potentially harmful event, but before/after an actual harm has occurred. The chief challenge here is to ensure *efficient ex post* enforcement. In order to achieve this, we focus on attaining maximum coverage of the region of enforcement, ensuring reliable and accurate detection of violation, and exploring a methodology to select *qualified* crowdsourced agents, called volunteers. We ensure maximum coverage of the given area of enforcement by proposing to divide it into regions using the Lloyd's algorithm and solving the enforcement problem by a divide and conquer mechanism over the entire area. We determine qualification of volunteers based on their likelihood of being in a region, over a given time interval and on trust, which is based on their *behavior* over the past. Finally, we use a non-incentive-based algorithm to select qualified volunteers for every region in the given area. We simulate the enforcement framework in CSIM19 (C++ version) and analyze the performance of the proposed volunteer selection algorithm over the area of enforcement.**

Keywords- volunteer; sentinel; ex post enforcement; crowdsourced spectrum monitoring; volunteer selection.

I. INTRODUCTION

With the exponential increase in use of wireless services, the demand for additional spectrum is steadily on the rise. In order to address this potential spectrum scarcity problem, the Federal Communications Commission (FCC) proposed Dynamic Spectrum Access (DSA), wherein licensed frequency bands when idle, are utilized by unlicensed users. In April 2015, the FCC adopted a three-tiered spectrum sharing infrastructure that is administered and enforced by Spectrum Access System (SAS) [1]. This architecture consists of Incumbents in tier 1, Priority Access Licensed (PAL) devices in tier 2 and General Authorized Access (GAA) devices in tier 3. Incumbents, in general, include military radars, fixed satellite service Earth stations and several of the Wireless Broadband Services (3650 – 3700 MHz) [2]. The SAS ensures that the spectrum is always available to the incumbent users when and where needed. The next level of access is provided to the users who buy PAL for a given location and period of time (usually for a three-year term). The remaining spectrum can then be used by devices having GAA. These devices have no protection from interference. They must, however, protect incumbents and PALs, while accessing spectrum [2].

As spectrum sharing becomes more intense and more granular with more stakeholders, we can expect an increasing number of potentially enforceable events. Thus, the success of spectrum sharing systems is dependent on our ability to automate their enforcement. The three key aspects of any enforcement regime are: the timing of enforcement action, the form of enforcement sanction and whether the enforcement action is private or public [3]. This paper focuses on detection of spectrum misuse. Thus, the key aspect of enforcement action for our consideration, is the timing of enforcement. Timing of an enforcement can be either *ex ante* (before a potentially "harmful" action has occurred) or *ex post* (after a potentially "harmful" action has occurred, but potentially before or after an actual "harm" has been done) [4]. The *ex ante* and *ex post* enforcement effects are inextricably linked. For example, if the *ex ante* rules and processes are sufficiently strong then *ex post* harms may be prevented before they occur. Also, certain types of *ex ante* rules may be easier to monitor and hence lower the cost of enforcement. Even strong *ex ante* rules may require *ex post* enforcement; for example, licensing approval for equipment is usually based on a prototype or pre-production unit, but compliance of production units may require some kind of policing. Till date, more significance has been given on automating *ex ante* enforcement of usage rights. As an example, the TV White Spaces database systems essentially

ISBN: 978-1-61208-695-8

work by preventing users with subordinate rights from using spectrum when and where other users with superior rights are operating [5]. This concept has been extended in the new Citizens Broadband Radio Service (CBRS) to a SAS that is designed to distinguish the three classes of user types discussed previously [2].

We observe that both SAS and CBRS have well-developed mechanisms to avoid interference but provide no support for addressing interference when it occurs. As we consider *ex post* enforcement approaches, the need to detect enforceable events, gather information about these events and adjudicate claims based on rules and evidence becomes important. In this paper, we focus on the detection of an interference event, or RF signal energy that is caused by a malicious user. The primary challenge is to ensure efficient *ex post* spectrum enforcement. In order to address this challenge, this paper proposes an enforcement framework that aims to achieve a) maximum coverage of the entire area of enforcement, b) an accurate, reliable and feasible detection of an event of violation, c) use of an effective method for hiring and deploying detecting agents. By employing a hybrid infrastructure of crowdsourced and trusted, dedicated resources, we aim to ensure “optimal” detection of spectrum access violation in Dynamic Spectrum Sharing Wireless networks. The major contributions of this paper are:

a) *Region Coverage*: We use a clustering algorithm to organize the area into smaller sized “regions” in order to ensure more manageable detection of violation
b) *Crowdsourced Detection*: We explore a mechanism to select crowdsourced detecting agents (called volunteers) for ensuring that a spectrum violation is detected with high probability of accuracy and efficiency.
c) *Volunteer Selection*: We develop a framework to assess the *qualification* of a volunteer across two dimensions - location likelihood and trust, to select volunteers using a non-incentive-based algorithm to ensure “optimal” quality of spectrum enforcement.

The paper is organized in the following manner. Section III of the paper discusses about the enforcement framework. Section IV discusses about the crowdsourced monitoring methodology, with a focus on the parameters that qualify a volunteer for selection and the appropriate volunteer selection mechanism. Section V discusses about the experimental setup and the results we obtained from applying the proposed volunteer selection algorithm. Finally, we conclude the paper and discuss about future works in Section VI.

II. Related Works

Jin *et al.* [20] introduces the first crowdsourced spectrum misuse detection for DSA systems. Dutta and Chiang [13] discusses about crowdsourced spectrum enforcement for accurate detection and location of spectrum enforcement. Salama *et al.* [22] proposed an optimal channel assignment framework for crowdsourced spectrum monitoring, where volunteers are assigned to monitor channels based on their availability patterns and are awarded with incentives in return. Li *et al.* [23] models the spectrum misuse problem as a combinatorial multi armed bandit problem to decide which channels to monitor, how long to monitor each channel, and the order in which channels should be monitored. Several incentive-based crowdsourced spectrum sensing works have been done over the past few years. Yang *et al.* [7] studied two incentive based crowdsourcing models, where a Stackelberg Equilibrium was computed in the platform-centric model, and a truthful auction mechanism was proposed under the user-centric model. Zhu *et al.* [14] proposes an incentive-based auction mechanism to improve fairness of bids by taking into consideration the effects of malicious competition behavior and the “free-riding” phenomenon in crowdsourcing services. Lin *et al.* [6] takes the Sybil attack into consideration for incentive based crowdsourced spectrum sensing. The works [11] and [12] propose frameworks for crowdsourced spectrum sensing without violating the location privacy of mobile users. Contrary to the formerly proposed spectrum monitoring approaches, which rely exclusively either on large deployment of physical monitoring infrastructure [8]-[10] or on crowdsourcing, we believe that spectrum misuse and access rights violations can be effectively prevented by using trusted infrastructure, composed of a central DSA Enforcement Infrastructure and a minimal number of mobile, wireless devices with advanced trust and authentication capabilities, augmented with an opportunistic infrastructure of wireless devices with various software and hardware capabilities. Moreover, in contrast to the usual methodologies, we explore the use of a non-incentive-based methodology for selection of volunteers to ensure maximum coverage of enforcement area and accurate detection of spectrum violations.

III. Enforcement Framework

The main challenge in the design of a hybrid infrastructure stems from the fact that it is not easy to determine where and how the resources are to be mobilized, given the non-deterministic nature of mobile devices’ *behavior*. It is equally difficult to determine how collaboration between these devices must take place to ensure swift detection and response to spectrum misuse and access rights violation. To address this, we broadly follow a crowdsourced monitoring infrastructure, supported by sentinel-based monitoring and a central DSA Enforcement Infrastructure.

A. System Model

The entire area of enforcement R is divided into smaller regions, with an Access Point AP_r, associated with every $r \epsilon R$. Authorized users, who are legitimate Secondary Users (SUs) gain access to an available channel through the local AP_r in r. On the contrary, malicious users are unauthorized transmitters who intrude on spectrum by illegitimately using spectrum frequencies in r that they have not been authorized to use by the local AP_r. Some of the authorized users volunteer to monitor a given channel for access violation, in addition to accessing the spectrum to transmit their own data. Such volunteers are mobile agents who can monitor radio access behavior within their neighborhood and detect anomalous use of spectrum. To carry out spectrum monitoring practices, volunteers incur transmit power consumption cost and bandwidth consumption cost.

As shown in Figure 1, the system model further consists of a central DSA Enforcement Infrastructure, which consists of a set of Volunteer Service units VS_r for every $r \in R$, a Volunteer Selection Unit and a DSA Database. A volunteer $v \epsilon V$ in $r \in R$ registers itself to the VS_r associated with r. A VS_r stores and updates volunteer attributes over the entire period of enforcement. The Volunteer Selection Unit uses the latest attributes of all the volunteers in a VS_r to select volunteers for monitoring a given channel in r over the next epoch of enforcement. The DSA Database maintains a channel-user occupancy list, for the entire area of enforcement R. The information contained in the DSA Database is used to identify the channels and their respective authorized users in R. Finally, the system model consists of a set of sentinels S' who monitor a given channel in r at random intervals to verify the detection results reported by the volunteers and to prevent selection of volunteers who have unreliable behavior.

B. Coverage of Region

To ensure maximum coverage of an area R for enforcement, we follow a divide and conquer method. We propose to divide the entire area R into smaller regions and then focus on solving the enforcement problem for a single region $r \in R$. This in turn can be used for solving the problem for the whole R. For division of R into regions, we propose the employment of the Voronoi algorithm [15]. Initially, we assume that the volunteers in V are randomly distributed over R and the access points are spread uniformly over R. For each volunteer $v \in V$, its corresponding Voronoi region r consists of every volunteer in the Euclidean plane whose distance to the local AP_r is less than or equal to its distance to any other AP_r [15]. However, the Voronoi algorithm may not produce regions that are of equal size. This is a disadvantage because it may result in some of the regions to have an undersupply of volunteers over time, which in turn may result in possible loss in detection of spectrum violation. Thus, we propose to apply a relaxation to the Voronoi algorithm, called the Lloyd's Algorithm [16], which produces uniformly sized convex regions, and thus improves the probability of a fair distribution of volunteers over all regions. The number of regions in R is equal to the number of access points in R.

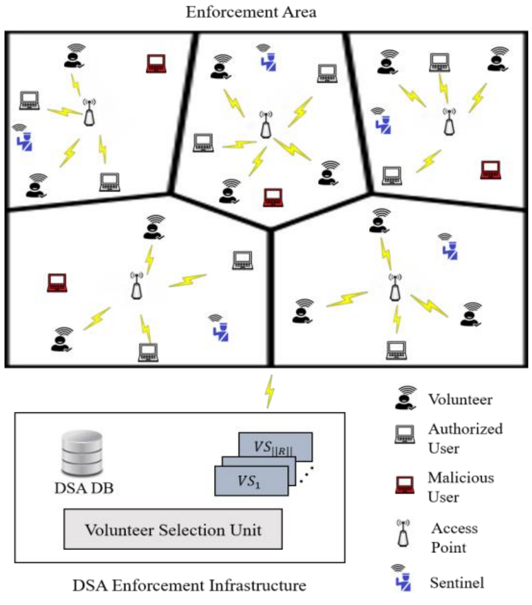


Figure 1. System Model.

IV. CROWDSOURCED SPECTRUM MONITORING

A volunteer $v \in V$ is associated with the following parameters: Serial Number of the sensing device S_v used by v and its location $L_{v,t}$ at time t. While S_v can be used to uniquely identify a volunteer, the location $L_{v,t}$ allows the VS_r of the DSA Enforcement Infrastructure to estimate whether v will be available to monitor a given channel in r in the future.

As shown in Figure 2, we divide the total enforcement time into a set of intervals called the Monitoring Intervals, MIs. Each MI is further divided into a set of n sub-intervals called the Access Unit Intervals (AUIs). One AUI is defined as the smallest interval over which a user, intruder or legitimate, can accomplish useful work. It is used as the interference monitoring interval by the selected volunteers to determine access violation or legitimacy. A new set of volunteers is selected at the end of every MI by the Volunteer Service unit VS_r associated with region r. Volunteer selection in r is primarily based upon twofold parameters of trust and location likelihood of a v in r.

A. Trust

The trust of a volunteer v is determined by its past behavior. The behavior of a volunteer v is chiefly determined by its accuracy in detection of spectrum violation. At the end of every AUI i, a volunteer v reports the observed state $\Phi_{i,v,r}$ of a channel c_r that it monitors, over i. The state of a channel c_r can be either a) violated, when c_r is being used by a malicious transmitter b) not violated, when c_r is either idle, i.e., when no user, authorized or malicious, uses c_r or safe, i.e., when c_r is used by an authorized transmitter. The necessary ground truth required for calculating accuracy of interference detection by v in r is acquired from the observed state $\Phi_{j,s,r}$ of c_r by a sentinel $s \in S'$ that monitors c_r at a random AUI j. A sentinel s is a trustworthy agent who helps in verifying volunteer detection result and helps to identify unreliable volunteers. As shown in Figure 2, a sentinel s monitors c_r in r at a random interval j,which is not known to the volunteers. This helps us to calculate the behavior $b_{i,v,r}$ of v in r at AUI i by using (1) given below.

$$b_{i,v,r} = \begin{cases} 1, & \Phi_{i,v,r} = \Phi_{j,s,r} \\ 0, & \Phi_{i,v,r} \neq \Phi_{j,s,r} \end{cases}, \forall i = j \tag{1}$$

As shown in (1), the behavior of a volunteer $b_{i,v,r}$ at i in r is assigned to zero when there is a mismatch in the observed state of c_r, between v and s.This can be because a) v makes a false detection, b) v lies about the true result, or c) s makes a false detection, d) s lies about the true result. However, for this paper, we assume that s is trustworthy and never makes a false detection or lies about a true result. An AUI when both v and s monitor channel c_r is called a matching interval. We aggregate $b_{i,v,r}$ over all the matching intervals to find the trust $T_{v,r}$ of v in r, by calculating the arithmetic mean $T_{v,r}$, given by (2),

$$T_{v,r} = \frac{1}{m}\sum_{p=1}^{m} b_{p,v,r} \tag{2}$$

where p is a matching interval and m is the total number of matching intervals over all the monitoring intervals observed so far.

B. Location Likelihood

In order to efficiently support detection of channel violation in a region r, volunteers who are most likely to reside a major proportion of time in r after a visit to r, must be given preference. For this purpose, the VS_r estimates the fraction of time that a volunteer v stays in r after its current visit to r. As shown in Figure 3, after the $(j)^{th}$ visit of v to r, we measure its $(j-1)^{th}$ sojourn time, $S_{j-1,v,r}$, in r as the difference between its $(j-1)^{th}$ departure time, $dep_{j-1,v,r}$ from r and its $(j-1)^{th}$ arrival time, $arr_{j-1,v,r}$ in r. Furthermore, we calculate the $(j-1)^{th}$ return time of v in r, $R_{j-1,v,r}$, as the difference between $arr_{j,v,r}$ and $arr_{j-1,v,r}$. As given by (3), this enables us to calculate the proportion of time, $P_{j-1,v,r}$, that v resided in r on its previous $((j-1)^{th})$ visit to r, as the ratio of $S_{j-1,v,r}$ to $R_{j-1,v,r}$. Based on this information, the VS_r estimates the proportion of time that v is likely to stay in r before its j^{th} departure from r, as an exponentially smoothed average, given by (4).

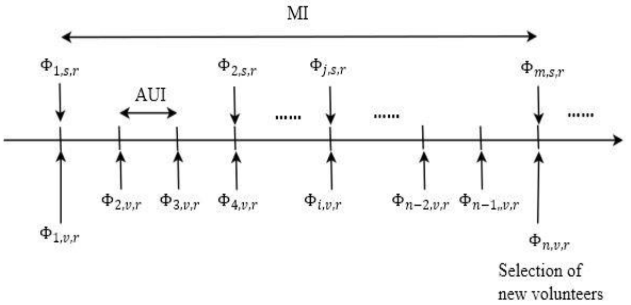


Figure 2. Observations $\Phi_{i,v,r}$ by volunteer v after every AUI and $\Phi_{j,s,r}$ by sentinel s after random AUIs, for the 1st MI.

$$P_{j-1,v,r} = \frac{S_{j-1,v,r}}{R_{j-1,v,r}} \tag{3}$$

$$\tilde{P}_{j,v,r} = \alpha . P_{j-1,v,r} + (1-\alpha) . \tilde{P}_{j-1,v,r} . \tag{4}$$

In order to estimate the smoothed average, $\tilde{P}_{j,v,r}$ more accurately, smoothing factor α is computed as:

$$\alpha = c\frac{E_{j-1,v,r}^{2}}{\sigma_{j,v,r}} . \tag{5}$$

where $0 < c < 1$, $E_{j-1,v,r} = P_{j-1,v,r} - \tilde{P}_{j-1,v,r}$ is the prediction error, and $\sigma_{j,v,r}$ is the average of the past square prediction errors on visit j. $\sigma_{j,v,r}$ can be expressed as follows:

$$\sigma_{j,v,r} = c . E_{j-1,v,r}^{2} + (1-c) . \sigma_{j-1,v,r} . \tag{6}$$

Moreover, at any given time t, the location $L_{v,t}$ of volunteer v enables us to estimate the likelihood of v to stay in r over the next monitoring interval, MI, based on the assumption that the likelihood of v to stay in r decreases as the displacement between $L_{v,t}$ and the centroid O_r of r increases. This is expressed by the separation factor, $\Upsilon_{t,v,r}$, given by (7) as follows:

$$\Upsilon_{t,v,r} = \gamma_1 e^{-\gamma_2 d(L_{v,t},O_r)} . \tag{7}$$

where $0 < \gamma_1, \gamma_2 < 1$, are parameters defined by the system and $d(L_{v,t}, O_r)$ is the displacement between $L_{v,t}$and O_r. Since $\Upsilon_{t,v,r}$ is exponential, so we empirically select values of γ_1 and γ_2 to avoid high variance in the values of $\Upsilon_{t,v,r}$ across all the volunteers.

Hence, the location likelihood, $L_{v,r}(MI)$ of v in r at time t over the next MI, is given by a function f of the parameters, $\tilde{P}_{j,v,r}$ of the latest (j^{th}) visit of v in r and $\Upsilon_{t,v,r}$,. We observe

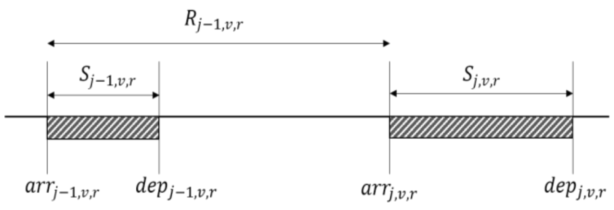


Figure 3. Sojourn time $S_{j,v,r}$ and Return time $R_{j,v,r}$ of volunteer v after its j^{th} visit to region r.

that since $R_{j-1,v,r} > S_{j-1,v,r}$ and $0 < \alpha < 1$, so $0 < \tilde{P}_{j,v,r} < 1$. Similarly, since $d(L_{v,t}, O_r) \geq 0$, so $0 < \Upsilon_{t,v,r} \leq 1$. As weighting the parameters by linear regression requires large amount of data and preferential weighting is hard to establish as it usually requires an expert opinion on the importance of an individual parameter relative to the overall composite parameter [17], so we assign equal weights to the parameters $\tilde{P}_{j,v,r}$ and $\Upsilon_{t,v,r}$. Finally, we define function f as the product of parameters $\tilde{P}_{j,v,r}$ and $\Upsilon_{t,v,r}$ as given by (8) below.

$$L_{v,r}(MI) = \tilde{P}_{j,v,r} \times \Upsilon_{t,v,r} \quad (8)$$

C. *Selection of volunteers*

From the set of volunteers, V, in total area of enforcement, R, the VS_r associated with r in the DSA Enforcement Infrastructure selects k_r qualified volunteers to monitor r at the beginning of every MI. This is determined by the estimated Qualification $Q_{v,r}(MI)$ of a volunteer v to monitor the associated channel c_r in r over the next MI, given by (9), defined below.

$$Q_{v,r}(MI) = g(T_{v,r},\ L_{v,r}(MI)) \quad (9)$$

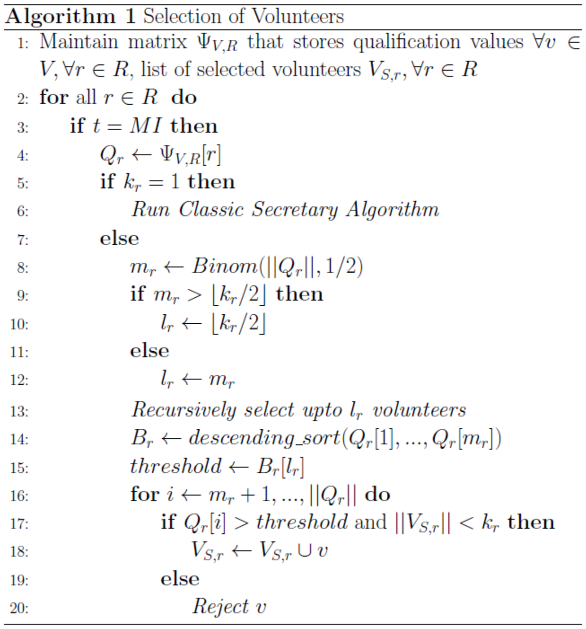
As shown in (9), $Qualification\ Q_{v,r}(MI)$ of a v in r is given as a function g of its location likelihood $L_{v,r}(MI)$, over the next MI, and trust $T_{v,r}$. Since $L_{v,r}(MI)$ and $T_{v,r}$ represent the measurement of two different parameters, we normalize both the parameters by using the min-max normalization technique [17]. We apply equal weighting to the two parameters since the other two widely used weighting methods of linear regression and preferential weighting are cumbersome due to the requirement of large amount of data and of expert opinion on preference, respectively [17]. We aggregate $L_{v,r}(MI)$ and $T_{v,r}$ in function g, using a) multiplication, b) addition, c) geometric mean, d) arithmetic mean and compare the performance of using different aggregation methods in Section V.

This work focuses on spectrum enforcement over a single channel in a region. However, it can be extended to deal with multiple channels in a region by selecting volunteers to cover additional channels. We also assume that a volunteer v can be hired to monitor more than one region over the next MI as v is mobile and can potentially cover multiple regions over a given MI. The Volunteer Selection Unit of the DSA Enforcement Infrastructure builds a centralized $||V||$-by-$||R||$ matrix $\Psi_{V,R}$, using the values of volunteer attributes from the VS_r associated with every region $r \in R$. The matrix $\Psi_{V,R}$ is a volunteer-region qualification matrix that contains the qualification values $Q_{v,r}(MI)$ of all $v \in V$ for every $r \in R$. The Volunteer Selection Unit selects k_r volunteers dynamically from V based on the qualification values of all $v \in V$ for r, using Algorithm 1.

For the volunteer selection Algorithm 1, we use the volunteer-region qualification matrix $\Psi_{V,R}$ to select qualified volunteers for every $r \in R$ (line 1). At the end of a MI (line 3), the Volunteer Selection Unit gains access to the qualification values of all $v \in V$ for r from $\Psi_{V,R}$ and stores them in a list Q_r (line 4). If the number of volunteers to be selected in r, k_r is 1, then we use the classic secretary algorithm [18] to select the most qualified volunteer dynamically, with constant probability. In a classic secretary algorithm, we observe the first $||Q_r||/e$ qualification values to determine a *threshold* and then select the first of the remaining volunteers, whose qualification value is above the threshold [19]. However, if $k_r > 1$, we select volunteers dynamically by using a variant of the multiple-choice secretary algorithm, which proceeds as follows. We draw a random sample m_r from a binomial distribution $Binomial(||Q_r||, \frac{1}{2})$, from which we select up to $\lfloor k_r/2 \rfloor$ volunteers recursively (lines 8-13). We keep appending the selected volunteers in set $V_{S,r}$. If m_r is greater than $\lfloor k_r/2 \rfloor$, then we set l_r to $\lfloor k_r/2 \rfloor$, otherwise we set l_r to m_r. Next, we set a $threshold$, which is the l_r^{th} largest qualification value in the sample of first m_r qualification values. After this, we select every volunteer with qualification value greater than $threshold$, till we select a maximum of k_r volunteers (lines 16-20) [19]. We apply this algorithm for selection of volunteers in every $r \in R$. The expected total qualification value of the k_r volunteers selected by Algorithm 1 is at least $(1 - \frac{5}{\sqrt{k_r}})$ times the total qualification value of the top k_r volunteers [19].

V. Experiments And Results

We simulate the enforcement framework by using the C++ version of the CSIM19 simulation engine. For simplicity, we



```
Algorithm 1 Selection of Volunteers
1: Maintain matrix $\Psi_{V,R}$ that stores qualification values $\forall v \in V, \forall r \in R$, list of selected volunteers $V_{S,r}, \forall r \in R$
2: for all $r \in R$ do
3:     if $t = MI$ then
4:         $Q_r \leftarrow \Psi_{V,R}[r]$
5:         if $k_r = 1$ then
6:             Run Classic Secretary Algorithm
7:         else
8:             $m_r \leftarrow Binom(||Q_r||, 1/2)$
9:             if $m_r > \lfloor k_r/2 \rfloor$ then
10:                $l_r \leftarrow \lfloor k_r/2 \rfloor$
11:            else
12:                $l_r \leftarrow m_r$
13:            Recursively select upto $l_r$ volunteers
14:            $B_r \leftarrow descending\_sort(Q_r[1], ..., Q_r[m_r])$
15:            $threshold \leftarrow B_r[l_r]$
16:            for $i \leftarrow m_r + 1, ..., ||Q_r||$ do
17:                if $Q_r[i] > threshold$ and $||V_{S,r}|| < k_r$ then
18:                    $V_{S,r} \leftarrow V_{S,r} \cup v$
19:                else
20:                    Reject $v$
```

Figure 4. Algorithm for selection of volunteers.

divide the entire area of enforcement R (of total area 500,000 sq. units) into two regions of equal area. This work can, however, be easily extended to deal with more regions. With the assumption that 1 sq. unit is equivalent to 1 sq. meter and by taking the average population density of Pittsburgh (2,140/sq. km) [21], we calculate the total population (1,070 people) in the area of enforcement. A random fraction of people from the total population are chosen as volunteers (equals 183 volunteers). Volunteers are initially placed at random positions within R and they change their positions with a random speed from the range 0-70 m/s at a random direction after every fixed interval of time. The maximum speed of a volunteer is chosen higher than the usual speed limit of a vehicle in a city in order to compensate for the limited simulation time. Additionally, we assume that every volunteer uses a sensing device with maximum battery capacity of approximately 7 Wh and that the battery discharges at the rate of 1 J/s for a random time interval drawn from an exponential distribution of the mean active time interval of 100 s. After every active time interval, we assume that the device remains idle for a random time interval drawn from an exponential distribution of the mean idle time interval of 10 s. The simulation runs till the battery of the sensing device used by every volunteer is exhausted, i.e., for 5665 AUIs. Each AUI is equivalent to 5 seconds and one MI is equivalent to 5 AUIs. We select $\gamma_1 = 1$ and $\gamma_2 = 0.01$ for the separation factor $\Upsilon_{t,v,r}$ of v with respect to r. Since $\Upsilon_{t,v,r}$ is exponential, so we empirically decide the value of the γ_2, which is the coefficient of $\mathrm{d}(L_{v,t}, O_r)$ from (7), to avoid high variances in the qualification values of volunteers. Since we did not notice significant difference in the results for different values of α, we determine the value of α empirically as 0.1 for the sake of simplicity in implementation. Finally, we assume that $k_r = k$ for every $r \in R$.

Primarily, we evaluate two performance metrics – the *hit ratio* and the *accuracy of detection*. In a monitoring interval MI, if a volunteer v selected for monitoring r is in r at the beginning of an AUI in the given MI, then it is a *hit*, otherwise it is a *miss* for the AUI in the given MI. This is according to the assumption that a selected volunteer v can successfully monitor a channel c_r in r over an AUI only if v resides in r over the given AUI. The *hit ratio* of a region $r \in R$ over a given MI measures the ratio of the number of *hits* of all the selected volunteers to the sum of the number of *hits* and the number of *misses* of all the selected volunteers in r. Figure 5 compares the mean *hit ratio* of all the regions over the entire duration of simulation, by using the proposed Algorithm 1 and the Random algorithm for different ranges of k. The Random algorithm selects k volunteers randomly from the total set of volunteers V. For this experiment, the qualification $Q_{v,r}(MI)$ of a volunteer v for region r is equal to its location likelihood $L_{v,r}(MI)$. We observe that Algorithm 1 has a better mean *hit ratio* than the random algorithm for all the ranges of k. However, the mean *hit ratio* by applying Algorithm 1 decreases consistently (from 0.91 for k =1-20% of $||V||$ to 0.57 for k = 20-40% of $||V||$) with the increase in k because the proportion of *qualified* selected volunteers reduces as the value of k increases. The error bars in Figure 5 represent the mean standard deviation of the mean *hit ratio* across all regions, which decreases from 0.22 for k =1-20% of $||V||$ to 0.101 for k = 80-100% of $||V||$, using Algorithm 1 and decreases from 0.19 for k =1-20% of $||V||$ to 0.06 for k = 80-100% of $||V||$, using the Random algorithm. This type of behavior is attributed to the fact that a balance is approached between the proportions of *qualified* and *unqualified* selected volunteers as the value of k increases.

We assume that every volunteer monitors the channel in r with a probability of successful detection p_v, drawn randomly from a uniform distribution in the range of $0 + \delta$ to $0.5 + \delta$ (with $\delta = 0.1$) and that a sentinel in r monitors the channel with the probability of successful detection of

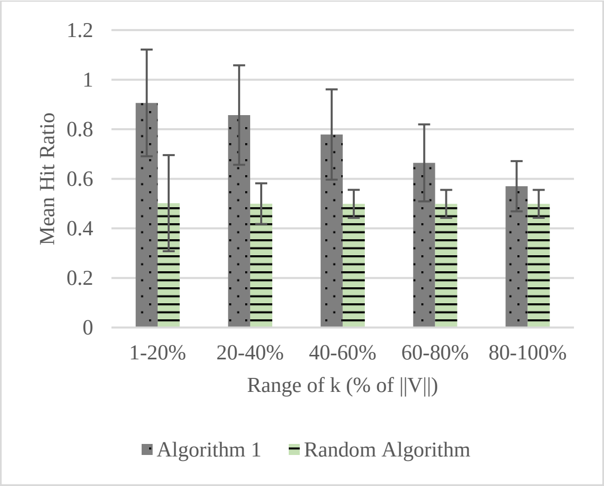


Figure 5. Comparison of the mean *hit ratio* of selecting volunteers by using Algorithm 1 and the Random algorithm.

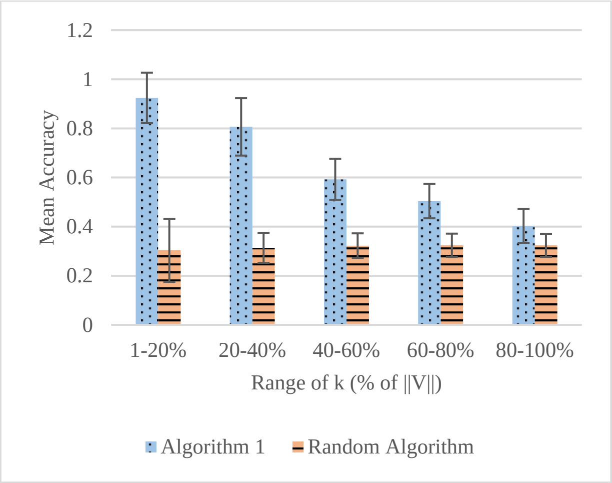


Figure 6. Comparison of the mean accuracy of detection by selecting volunteers using Algorithm 1 and Random algorithm.

0.5. The observed detection value of a v at the end of an AUI is a random sample taken from a binomial distribution of 100 trials with a probability of success equal to p_v. Similarly, the observed detection value of a sentinel $s \in S'$ at the end of an AUI is a random sample from a binomial distribution of 100 trials with a probability of success equal to 0.5. If the observed detection values of a v and s in r are both either greater than or lesser than an empirically selected threshold value of 42, then the detection by v is considered accurate. Figure 6 compares the mean accuracy of detection of the selected volunteers over all the MIs between Algorithm 1 and the Random algorithm for varying ranges of k. For this experiment, the qualification $Q_{v,r}(MI)$ of a volunteer v for region r is equal to its trust $T_{v,r}$. We observe that Algorithm 1 performs better than the Random algorithm for all the ranges of k. The mean accuracy of detection decreases consistently (from 0.92 for k = 1-20% of $||V||$ to 0.403 for k = 80-100% of $||V||$) with the increase in k because of the decrease in the fraction of *qualified* volunteers in r as k increases. The uniform distribution of p_v for all $v \epsilon V$ ensures that the mean standard deviation in accuracy of detection across all regions decreases from 0.103 for k =1-20% of $||V||$ to 0.069 for k = 80-100% of $||V||$, using Algorithm 1 and decreases from 0.13 for k =1-20% of $||V||$ to 0.05 for k = 80-100% of $||V||$, using the Random algorithm.

Finally, in Figure 7, we compare the mean *hit ratio* and the mean accuracy of detection by using different data aggregation methods to aggregate the trust $T_{v,r}$ and location likelihood $L_{v,r}(MI)$ of v to calculate it's qualification $Q_{v,r}(MI)$ using (9) for region r over the next MI. For this experiment, we select volunteers using Algorithm 1 for k = 1-20% of $||V||$. We observe that by using the aggregation methods of sum and arithmetic mean, we obtain a mean accuracy value (equals 0.86) higher than the mean *hit ratio* (equals 0.76). On the contrary, we observe that the mean accuracy value (equals 0.78) is lower than the mean *hit ratio* (equals 0.85) when we use the data aggregation methods of product and geometric mean. Also, we observe that the results we obtain by using sum and arithmetic mean are similar. Likewise, we get similar results for both product and geometric mean. This would help us to decide about the proper aggregation method to use based on the requirements of the system for efficient spectrum enforcement.

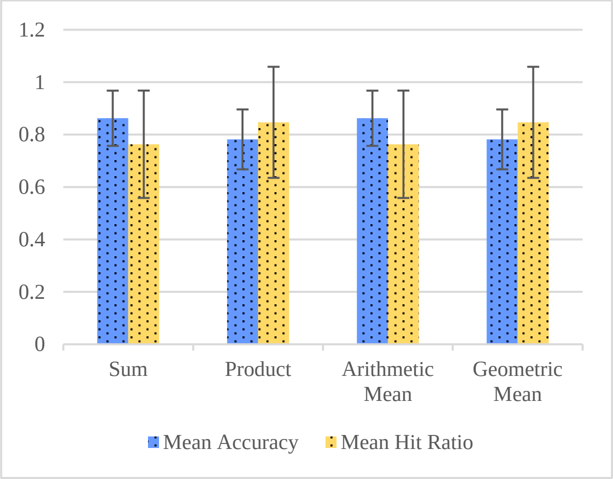


Figure 7. Comparison of the mean hit ratio and mean accuracy of detection by selecting volunteers using Algorithm 1 with different aggregation methods to calculate $Q_{v,r}(MI)$ of volunteer v in r for k = 1 to 20% of $||V||$.

VI. Conclusion

In this paper, we discuss about a spectrum enforcement framework based on a crowdsourced monitoring infrastructure, supported by sentinel-based monitoring and a central DSA Enforcement Infrastructure. The objective is to maximize coverage of the area of enforcement and to ensure reliable detection of spectrum access violation by selecting *qualified* volunteers. We propose to maximize the coverage of the region of enforcement by following a divide-and-conquer mechanism wherein we divide the area of enforcement into smaller regions, by applying the Lloyd's algorithm, which is a relaxation to the Voronoi algorithm. Every small region in the enforcement area is responsible for its own spectrum enforcement, which in turn ensures enforcement of the entire area. The qualification of a volunteer for the upcoming time interval is decided by its likelihood to stay in the region over the next monitoring interval and by its trust. We use a variant of the multiple-choice Secretary algorithm to select volunteers dynamically based on their qualifications to monitor a region. We observe that this non-incentive-based volunteer selection algorithm performs better than a non-incentive-based algorithm that selects volunteers randomly for spectrum monitoring.

We plan to extend this work to explore different mechanisms to select volunteers for multi-channel spectrum enforcement. We further plan to explore different statistical and machine learning based mechanisms to determine the trust and location likelihood of volunteers in the enforcement area.

Acknowledgment

This work was sponsored in part by the National Science Foundation through grants 1265886, 1547241, 1563832, and 1642928.

References

[1] Federated Wireless. *Citizens Broadband Radio Service (CBRS) Shared Spectrum: An Overview*. [Online]. Available from: http://federatedwireless.com/wp-content/uploads/2017/03/CBRS-Spectrum-Sharing-Overview-v3.pdf.

[2] Federal Communications Commission. *3.5 GHz Band / Citizens Broadband Radio Service*. [Online]. Available from: https://www.fcc.gov/wireless/bureau-divisions/broadband-division/35-ghz-band/35-ghz-band-citizens-broadband-radio#block-menu-block-4.

[3] E. Schlager and E. Ostrom, "Property-Rights Regimes and Natural Resources: A Conceptual Analysis," Land Econ., vol. 68, no. 3, 1992, pp. 249–262.

[4] Shavell, Steven. “The Optimal Structure of Law Enforcement.” The Journal of Law & Economics, vol. 36, no. 1, 1993, pp. 255–287. JSTOR, www.jstor.org/stable/725476.

[5] A. Gopinathan, Z. Li, and C. Wu, “Strategyproof auctions for balancing social welfare and fairness in secondary spectrum markets,” 2011 Proc. IEEE INFOCOM, 2011, pp. 3020–3028.

[6] J. Lin, M. Li, D. Yang, G. Xue, and J. Tang, “Sybil-proof incentive mechanisms for crowdsensing,” in IEEE INFOCOM 2017, pp. 1–9.

[7] D. Yang, G. Xue, X. Fang, and J. Tang, “Crowdsourcing to Smartphones: Incentive Mechanism Design for Mobile Phone Sensing,” in Proceedings of the 18th Annual International Conference on Mobile Computing and Networking, 2012, pp. 173–184.

[8] M. B. H. Weiss, M. Altamimi, and M. McHenry, “Enforcement and spectrum sharing: A case study of the 1695-1710 MHz band,” in 8th International Conference on Cognitive Radio Oriented Wireless Networks, 2013, pp. 7–12.

[9] D. Yang, X. Zhang, and G. Xue, “PROMISE: A framework for truthful and profit maximizing spectrum double auctions,” in Proceedings - IEEE INFOCOM, 2014, pp. 109–117.

[10] R. Chen, J.-M. Park, and J. H. Reed, “Defense Against Primary User Emulation Attacks in Cognitive Radio Networks,” IEEE J.Sel. A. Commun., vol. 26, no. 1, Jan. 2008, pp. 25–37.

[11] X. Jin, R. Zhang, Y. Chen, T. Li, and Y. Zhang, “DPSense: Differentially Private Crowdsourced Spectrum Sensing,” in Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, 2016, pp. 296–307.

[12] X. Jin and Y. Zhang, “Privacy-Preserving Crowdsourced Spectrum Sensing,” IEEE/ACM Trans. Netw., vol. 26, no. 3, Jun. 2018, pp. 1236–1249.

[13] A. Dutta and M. Chiang, ““See Something, Say Something” Crowdsourced Enforcement of Spectrum Policies,” IEEE Trans. Wirel. Commun., vol. 15, no. 1, Jan. 2016, pp. 67–80.

[14] X. Zhu, J. An, M. Yang, L. Xiang, Q. Yang, and X. Gui, “A Fair Incentive Mechanism for Crowdsourcing in Crowd Sensing,” IEEE Internet Things J., vol. 3, no. 6, Dec. 2016, pp. 1364–1372.

[15] F. Aurenhammer, “Voronoi diagrams—a survey of a fundamental geometric data structure”, ACM Comput. Surv., vol. 23, no. 3, Sep. 1991, pp. 345–405.

[16] Q. Du, M. Emelianenko, and L. Ju, “Convergence of the Lloyd Algorithm for Computing Centroidal Voronoi Tessellations,” SIAM J. Numer. Anal., vol. 44, no. 1, Jan. 2006, pp. 102–119.

[17] B. Talukder, K. W. Hipel, and G. W. vanLoon, “Developing Composite Indicators for Agricultural Sustainability Assessment: Effect of Normalization and Aggregation Techniques,” Resources, vol. 6, no. 4, 2017.

[18] Gautam Kamath. *Advanced Algorithms, Matroid Secretary Problems.* [Online]. Available from: http://www.gautamkamath.com/writings/matroidsec.pdf.

[19] R. Kleinberg, “A Multiple-choice Secretary Algorithm with Applications to Online Auctions,” in Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, 2005, pp. 630–631.

[20] X. Jin, J. Sun, R. Zhang, Y. Zhang, and C. Zhang, “SpecGuard: Spectrum misuse detection in dynamic spectrum access systems,” 2015 IEEE Conf. Comput. Commun., 2015, pp. 172–180.

[21] Pittsburgh Population. (2018-06-12). [Online]. Available from: http://worldpopulationreview.com/us-cities/pittsburgh/.

[22] A. M. Salama, M. Li, and D. Yang, “Optimal Crowdsourced Channel Monitoring in Cognitive Radio Networks,” in IEEE Global Communications Conference, GLOBECOM, Singapore, December 4-8, 2017, pp. 1–6.

[23] M. Li, D. Yang, J. Lin, M. Li, and J. Tang, “SpecWatch: A framework for adversarial spectrum monitoring with unknown statistics,” Comput. Networks, vol. 143, 2018, pp. 176–190.