Sustenance against RL-based Sybil attacks in Cognitive Radio Networks using Dynamic Reputation Systems

Kenneth Ezirim, Erald Troja
Department of Computer Science
Graduate Center, CUNY
New York, NY 10016
Email: {kezirim,etroja}@gc.cuny.edu

Shamik Sengupta
Department of Math. & Comp.Sc.
John Jay College, CUNY
New York, NY 10019
Email: ssengupta@jjay.cuny.edu

***Abstract*—Sybil attacks are known form of denial-of-service attacks that are common-place in Dynamic Spectrum Access networks. In this paper, we formulate novel threat and defense mechanisms for the Sybil attack problem in Cognitive Radio Networks (CRN). We present potential identity sampling strategies that a malicious Sybil attacker can use to enhance its attack capability and impact without being detected. We investigate how a Sybil attacker can leverage reinforced learning to improve its performance. We also formulate a novel dynamic reputation mechanism to defend against such threat that relies on the nodes' reporting in an intelligent and adaptive manner. Results obtained shows that a Sybil attacker can improve its performance using RL learning technique. It also demonstrates that the use of the dynamic reputation mechanism can considerably reduces the effectiveness of Sybil attacks and improve the accuracy of spectrum decisions.**

I. Introduction

Dynamic Spectrum Access (DSA) offered through cognitive radios (CR) is one of the most promising frameworks for alleviating the crunching pressure put forth on the FCC by many of the wireless providers. In contrast to the legacy fixed spectrum allocation policies, DSA allows license-exempt secondary-users (SUs) to access the licensed spectrum bands when not in use by the licensed owners, also known as primary users (PUs). DSA is expected to enable more efficient use of frequency channels without impacting the primary licensees [1].

One of the major challenges to DSA's success is the provision of security for CR networks. The "open access" philosophy of the FCC-proposed DSA paradigm makes cognitive radio networks susceptible to various unforeseen attacks by smart malicious societies. The effect of malicious disruptions can be even more fatal as there is no way to understand whether the disruptions are unintentional or intentional. The motivation for such *shadow-disruptive attack behavior* can be either monopolism – to capture as much spectrum as possible for themselves without maintaining any spectrum sharing etiquettes and make other secondaries starve [2]; or adversarial – to disrupt other secondaries' communications and shut them down (particularly applicable in environments filled with adversary users/networks) [3], [4]. The current DSA paradigm lacks the protocols to handle most of the security issues arising in a CR network [5]. The Sybil attack, which is classified as an identity spoofing attack, is one of the most common security issues in CR network. During a Sybil attack, a malicious attacker operates with multiple identities, pretending to be multiple distinct entities [6]. The sybil threats using malicious cognitive radio(s) in DSA networks are even more prevalent and dangerous for several reasons:

1) They are highly "mobile" in every possible aspect due to the characteristics of software reconfigurability;
2) DSA networks are susceptible to attacks ranging from passive eavesdropping to active interfering, frequent break-ins by adversaries due to their open, ubiquitous and interoperable nature [7];
3) Due to the open source nature of DSA networks, it is practically impossible to establish a standard database to record the identity information of every CR node [8].

In this paper, we explore the dynamic behavior of the Sybil attacker which operates in a CR network where CR nodes conduct spectrum sensing in a local and distributed manner and report the results to a centralized Fusion Center (FC) which in turn take decision of spectrum availability based on all the spectrum usage reports. Each node has a unique identity associated with every sensing report it makes to the FC about the spectrum state. The attacker is capable of generating multiple false sensing reports directed at the FC with the help of its multiple Sybil identities. We implement a number of identity sampling strategies that an attacker can use to select the number of identities k to be used for false reporting. Based on feedback from the FC and with the help of reinforced learning (RL), the Sybil dynamically adjusts k for the next stage of attack.

As a defense mechanism against Sybil attacks, we introduce an adaptive Fusion Center that implements a dynamic reputation mechanism for spectrum decision making. The dynamic reputation mechanism uses a non-linear reputation function to compute the reputation of an identity, which reflects to a great extent the integrity of its sensing reports.

The rest of this paper is organized as follows: Section II presents the system model of a potential Sybil attacker and FC. In section III, we discuss about the Sybil attacker, the learning strategy, identity ranking algorithm and identity sampling strategies used in attacks. Section IV is dedicated to the FC's dynamic reputation system. Section V shows the numerical and simulation results obtained from our experiments. Conclusions are finally drawn in the last section of this paper.

II. System Model

We consider a CR network, comprising of N honest secondary nodes and malicious Sybil attacker(s). The malicious node as well as the honest secondary nodes have the capability of sensing the spectrum periodically and reporting their results to the FC. Sensing reports indicates either primary user's absence or presence in the spectrum, which can be interpreted as 0 or 1 respectively.

After reporting, all nodes await FC's decision before commencing activity in the spectrum. The aim of the Sybil attacker is to maximally corrupt sensing result, thereby forcing the FC into wrong conclusions. Wrong conclusions drawn by the FC are costly and leads to either under-exploitation of the spectrum or conflict with the primary user. The attacker derives its utility by successfully causing a disruption in the activities of the honest nodes. In this model, FC is capable of verifying its conclusions and computing reputation of identities based on its verification results. However, there is an error probability associated with FC's verifications, which we refer to as verification error probability ξ. The verification process can be conducted with the help of policy nodes [9] or simply by "listening" on the channels to detect contentions or white spaces.

III. The Sybil Attacker

We assume that Sybil attacker can generate up to M identities. For example, in our experimental testbed setup, Soekris Net-5501 boards enabled with Atheros chipsets based Wireless NIC and programmable Madwifi device driver, we are able to support up to 64 identities for one physical device [2]. The attacker is aware of FC's vigilance on relegating suspicious identities by assigning low reputation to those identities. At every stage of reporting, it uses $k \in [0 \cdots M]$ to report falsely and $M-k$ to report truly the spectrum state. A greedy attacker would use $k = M$ identities to report falsely to the fusion center but this approach will expose the Sybil identities to the fusion center in a very short time. The best strategy to avoid detection would be to find the optima k that would have maximum impact on the fusion center decision. Since the main objective of the attacker is to avoid detection which reduces its reputation, it should selectively choose the k identities used for the attack. The selection is done such that the frequency of use of each identity in reporting to the FC is kept under control. Later on, we shall consider various identity sampling strategies including the identity hopping strategy used by the attacker.

A. Sybil Attacker Learning Strategy

The Sybil attacker's goal is to maximize the effectiveness of its attacks (i.e. its performance) on the FC. The attacks have to be unpredictable and in such a manner that the Sybil identities are not compromised. Given a system of CR networks, characterized by several characteristics(parameters) such as arrival rate of primary users, sensing error probability of CRs and verification error probability of the FC, the Sybil attacker can learn the system to know the optimal strategy or policy to implement in order to corrupt FC's decision.

In considering the optimality of k, the Sybil attacker has to consider its level of success in the previous stages of attacks. The knowledge of its previous attack strategies that were successful will enhance the decision making process of the attacker in ascertaining the right number of identities to use in embarking on an attack. It must be noted that since the cumulative number of successful attacks varies in each time slot, k needs to be kept dynamic. Another reason why k should be varied is to accommodate the mechanism that allow identities to recover their reputation after a period of intensive attacks. Keeping k constant could reveal the periodicity or the distribution used in sampling the identities.

Given the complexity and dynamism of the wireless environment, reinforcement learning (RL) can provide the Sybil attacker the context awareness and intelligence to conduct its attack efficiently in the system. RL have been applied in many aspect of wireless network such as routing, resource management and dynamic channel selection. We adopt the Q-learning algorithm in learning the Sybil attacker on the best k to use at any level of performance measure.

When making false reports with k of its identities, the goal of an attacker is to get the FC agree with whatever those identities reported. Let $\mathcal{Z}(t) \in \{0,1\}$ be an indicator function of time that shows whether FC concluded in favor of the attacker. While using the k identities, the attacker risks exposing them. The attacker associates a cost $\mathcal{K}(t)/M$ with the exposure, where $\mathcal{K}(t)$ is a function that returns the value of k at time t. An attacker can have a variable belief $\mu \in [0,1]$ which represents the level of it conviction that the identities will be penalized by the FC for false reporting. So after t_c stages of attacks, attacker's perceived performance $P(t_c) \in [0,1]$ is expressed as follows

$$P(t_c) = \frac{1}{M \cdot t_c} \cdot \sum_{t=1}^{t_c} [M \cdot \mathcal{Z}(t) - \mu \cdot \mathcal{K}(t)] \qquad (1)$$

In a case where $\mu = 0$, which indicates attacker's utmost confidence that its identities would not be discovered by FC, equation 1 simplifies to

$$P(t_c) = \frac{1}{t_c} \cdot \sum_{t=1}^{t_c} \mathcal{Z}(t) \qquad (2)$$

With $\mu = 1$, which indicates that the attacker is certain of the risk of exposure, we can observe the dependency of $P(t_c)$ on the value of k used at each stage of the attack. There is no gain for the attacker using $k = M$ in all the stages of the attack, even if all attacks were successful. So it is not in the best interest of the attacker to use all its identities but rather intelligently choose k of best fit for an attack.

For the purpose of implementing RL based Sybil attack, we need to clearly define the terms: state and action. We define a state s_t of an attacker as an indicator of its performance level at time t. The attacker defines an arbitrary finite number of states η solely to differentiate its performances level and learn the best action to take at each level. As result, the range of $P(t_c) \in [0,1]$ is partitioned into η possible intervals and each interval is associated with a unique state. Assuming $\eta = 4$ and $P(t_c)$ is uniformly partitioned, then the state s_t of an attacker can be determined using the following rules:

$$s_t = \begin{cases} 1 & 0.0 \le P(t) < 0.25 \\ 2 & 0.25 \le P(t) < 0.5 \\ 3 & 0.5 \le P(t) < 0.75 \\ 4 & 0.75 \le P(t) < 1.0 \end{cases}$$

Worst performance of attacker is therefore associated with state $s_t = 1$, and the best performance, with the state $s_t = 4$.

For a uniformly partitioned interval, s_t can be expressed concisely as

$$s_t = \lceil P(t) \cdot \eta \rceil,\ s_t \in \mathbb{Z}_{>0} \tag{3}$$

An action, however, is defined by the number of identities k that is used by attacker in making false report to FC. For example in state $s_t = 3$, an attacker can choose to use $k = 4$ identities, where $k \in [1, M]$, to make false report. Using Q-learning, some other important parameters include learning rate denoted by α and discount factor denoted by γ. The reward r is dependent on the success of an attack and k. In the case of success the reward assigned is inversely proportional to the k but in case of failure, the reward is negative and directly proportional to k. The reward is designed in such a manner to discourage any tendency to "greedy attack" i.e. $k \rightarrow M$. The Q-learning algorithm uses the function $Q : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{R}$ to compute the quality of a state $s \in \mathcal{S}$ after performing an action $a \in \mathcal{A}$. $\mathcal{S}$ is a set of possible of states and $\mathcal{A}$ is a set of possible actions. After carrying out an attack at time t, the attacker updates the value of the state-action pair (s_t, a_t) in its Q-table as follows:

$$\begin{aligned} &Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) \\ &+ \alpha \cdot \left[r(a_t) + \gamma \cdot \max_{a_{t+1}} Q_t(s_{t+1}, a_{t+1}) - Q_t(s_t, a_t) \right] \end{aligned} \tag{4}$$

where s_{t+1} and a_{t+1} are the new state (after performing action a_t) and the best action in the new state respectively. The iterative nature of the algorithm uncovers the best action to be taken given a certain state of performance. The best action a^*, given the current state s_t, is the action $a \in \mathcal{A}$ with the maximum Q-value which is given by the equation

$$a^* = \max_{a \in \mathcal{A}} Q(s_t, a) \tag{5}$$

B. Identity Ranking Algorithm

An identity i used by the Sybil attacker is characterized by the number of successful attacks z_i and the number of exposures e_i. The attackers in order to maintain a balance between the effectiveness of its attacks and its reputation with the FC has to select identities that have been less exposed but with high success per exposure. So the first criterion ensures that the least exposed identities are used in attack. The second criterion further ensures that the least exposed identities have been effective in previous attacks.

The attacker can use the following algorithm to rank the identities, before selecting the best k candidate identities for an attack. Let us assume that the attacker has a set of identities $\mathcal{I} = \{(z_i, e_i) | i \in M\}$ to rank. We can decide to use a sorting algorithm (merge sort) with an optimal complexity of $O(nlogn)$ to rank the identities in ascending order. In that case the merge function is modified as follows. Two identities i and j are compared based on the ratios z_i/e_i and z_j/e_j. The identity with the least ratio is appended first. If the ratios of the identities are equal, then e_i and e_j are compared. Since e_i and e_j are measures of exposure, the identity with the least exposure is appended first. It is obvious that the complexity algorithm is not affected by the modification. The outcome of the modified sorting algorithm is a set $\mathcal{I}^*$ of ordered (ranked) identities.

C. Identity Sampling Strategies

The identity sampling strategy adopted by the Sybil attacker is targeted towards evading discovery by the FC. So the attacker can choose to sample its identities in a particular order that reduces the frequency of their appearance in attacks. Some of the strategies that we considered are as follows: sliding window, best performing k, and identity hopping.

1) Sliding Window Strategy

The sliding window (SW) strategy selects identities based on specified window, whose size is controlled by k. The identities are arranged in such an order that they form a loop (see Figure). In each round, the window slides fixed number of position(s) to the right. The identities that falls within the window are used to conduct the attack. The identities are arranged to form a loop such that the window simply wraps around. As we can see, the strategy ensures that in every round, one new identity is featured and one old identity is withdrawn. The sliding window strategy exhibits periodicity which is dependent on the number of positions it slides across. As a result, the identities are periodically sampled for the attack. An illustration of the sliding window identity sampling strategy can be seen in Figure 1.

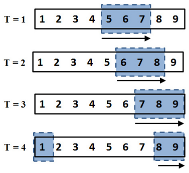

Fig. 1. Sliding Window Identity Sampling Strategy

2) Best Performing k Strategy

In best performing k (BP) strategy, the attacker samples those identities obtained from the identity ranking algorithm described. The criteria of the algorithm, which favors identities with relatively better performance with minimum exposure, ensures no identity with high failure is repeated frequently. The best performing k strategies leverages this approach to inject some sort of balance between success of an attack and the reputation of identities.

3) Identity Hopping Strategy

Inspired by the concept of frequency hopping mechanism [10], the identity hopping algorithm is targeted towards generating a pseudo-random sequence of identities. It involves ranking of the identities, which is carried out using the IR algorithm described. The ranking process is followed by the generation of a pseudo-random sequence which relies on ranking results obtained. The sequence generated allows some bias towards identities with higher ranking. Using a decreasing linear function, the number of occurrence y_i of each identity in the sequence is determined using the expression: $y_i = M - r_i + 1$, where r_i is the identity's rank. A decreasing parabolic function can also be used instead, defined as $y_i = (M - r_i + 1)^2$. In this paper, we assume that linear function is used in the pseudo-random sequence generation. Consider the following example of the pseudo-random sequence generation. Assuming that the identities are ranked as follows $\{1, 4, 2, 5, 3\}$, then identity 1 occurs 5 times in the sequence, 4 occurs 4 times, 2 occurs 3 times etc.

The k identities used in the attack are selected from the sequence with replacement. It is obvious that the probability of selecting the highest ranked identity is high given that the length of the sequence is $M(M+1)/2$. But we cannot ignore that it is still probable to pick an identity other than highest ranked identity. The randomness associated with this strategy makes it difficult for the FC to learn the sampling pattern of the Sybil attacker. The attacker achieves a dual goal here: increasing the chances of selecting the best performing k identities and randomizing its choices.

IV. The Fusion Center and Reputation System

Most reputation mechanisms are implemented using a linear reputation function that is dependent on nodes' performance [9]. The problem with linear reputation system is that it penalizes every node equally, even the honest nodes. There is no mechanism to deal with each node separately based on its performance.

Here, we introduce a reputation mechanism that relies on a non-linear reputation function to compute the reputation of nodes. The reputation function is a non-decreasing exponential function that converges to 1. It implies that using this reputation mechanism, reputation values of identities can only vary between 0 to 1. The reputation value is assigned by a reputation function $r_i(x,m)$ that is given by the following expression:

$$r_i(x,m) = \left[\sum_{j=0}^{m} \frac{x^j}{j!}\right] / e^x \tag{6}$$

where x is the parameter that controls the shape of the function and m is the parameter that indicates the number of correct reports made by using i-th identity. Depending on the value x, the reputation function converges quickly or slowly to value of 1. Increasing x leads to a slower convergence and decreasing x leads to a faster convergence. The parameter m is a measure of performance of an identity. It is important in the computation of the reputation of an identity and control of the reputation function's dynamics. The parameter m is updated as follows:

$$m = \begin{cases} m+1 & \text{correct report} \\ m-1 & \text{false report} \end{cases} \tag{7}$$

The dynamics of the reputation function is illustrated in Figure 2. We introduce a new parameter δx that determines by how much x is updated. By increasing x by δx, the shape of the reputation function changes. Using Figure 2 and assuming $x = 7$, an increase by $\delta x = 3$ makes the shape of the reputation function to become more concave (with respect to x-axis), indicating slower growth of reputation with increasing m. Likewise, a reduction in x by $\delta x = 2$, assuming $x = 2$ makes the shape of the function to be more convex, indicating faster growth of reputation with continued correct reporting.

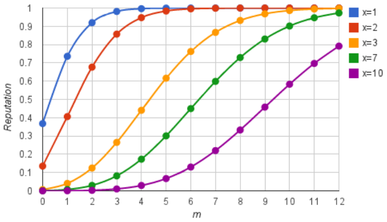

Fig. 2. Reputation function curves with different values of x.

The fusion center can implement a strategy such that new nodes are assigned a slow converging reputation function. Based on a node's performance over time, its reputation function would be tuned to allow for better growth rate of its reputation with the fusion center. A node is also penalized for reporting incorrectly by assigning it a reputation function that has a slow reputation growth rate. On the other hand, if a node's performance is satisfactory, it will be rewarded by improving its reputation function with a faster growth rate.

Depending on fusion center's priorities a tolerance threshold can be determined for both good and bad reporting. The tolerance threshold for good reporting τ_G determines when x should be decreased to improve the reputation function. For an FC with high standard, τ_G would be set high enough to ensure that a node's overall performance has significantly improved before its x is reduced. The same applies in the case of determining the tolerance threshold for bad reporting τ_B. A strict FC would chose a significantly low τ_B to tolerate fewer cases of false reporting. Therefore, $\tau_B = 0$ would indicate zero tolerance for false reporting.

The fusion center decides on primary user presence or absence using the computed reputation values of identities. Each identity makes a report $\theta \in \{0,1\}$ to fusion center. Based on θ the identities are separated into two groups: group G_0 where $\theta = 0$ and group G_1, where $\theta = 1$. The reputation values of members of a group are added up. The reputation $r_\theta(t)$ of identities belonging to G_θ at time t is given as

$$r_\theta(t) = \sum_{i}^{|G_\theta|} r_{\theta,i}(t) \tag{8}$$

where $i \in G_\theta$ is the i-th identity that reported θ. Based on the accumulated reputation the fusion reaches the final decision. The final decision θ^* is given by the expression:

$$\theta^* = \underset{\theta\in\{0,1\}}{\operatorname{argmax}}\{r_\theta(t)\} \tag{9}$$

In a case where both groups' cumulative reputations are equal or sums up to zero reputation, the majority rule is used in breaking the tie.

V. Numerical and Experimental Results

We conducted simulation experiments to evaluate the performances of the RL based Sybil attacker and the adaptive Fusion Center. In all experimental setups, we focused on scenarios where $M \geq N$, that is the number of Sybil identities exceeded or is equivalent to the number of honest identities. We assume for the sybil attack implementation using RL, $\gamma = 0.1$ and $\alpha = 0.8$. We kept the number of possible states and actions respectively as $\eta = |\mathcal{S}| = 10$ and $|\mathcal{A}| = M$. For the FC, we assumed that $\tau_B = 0.2$, $\tau_G = 0.85$ and $\delta x = 0.05$ for simulation purpose.

The main intent behind the experiments is to underscore the importance of RL in the effectiveness of Sybil attacks and show that FC is capable of improving the accuracy of its decisions by implementing a dynamic reputation mechanism (DRM). The performance of non-linear DRM is compared against the static linear reputation mechanism (SRM) mentioned in [9]. We also intend to make a comparison of the performances of Sybil identity sampling strategies that could be used in a Sybil attack to mitigate exposure.

A. *Performance of RL based Sybil Attack*

In this paper, we discussed how the Sybil attacker can leverage the RL technique to determine the best k to use in attacking the FC. We carried out some experiments to determine how the use of RL can enhance the performance of the Sybil attacker. To do this, we considered two different methods employed by FC in making decision about spectrum availability: static reputation mechanism (SRM) and dynamic reputation mechanism (DRM). In order to demonstrate the benefit of employing RL in the attack, we further embark on comparing the performances of a RL based attack and a random attack. Random attack involves the attacker randomly selecting the number of identities k to use for false reporting in each stage of the attacks. It is comparable to an RL based attack with $\epsilon = 1$ i.e. 100% exploration.

Using the SRM mechanism, we consider the following cases where an attacker implements: random attack, RL based attack with $\epsilon = 0.3$ and RL based attack with $\epsilon = 0.8$. The results obtained for the cases just stated are illustrated in Figure3. In general, we observe that the effectiveness of the Sybil attacks improves with increased verification error ξ of the FC. This trend is expected as an increase in ξ indicates that FC cannot correctly identify the false-reporting identities and penalize them. The figure further reveals the benefits of using an RL technique (Q-learning) in conducting Sybil attacks. It can be seen that it is more effective if RL technique is employed in the attacks with small ϵ. In Figure 3, it is obvious that best performance was mostly attained with varying ξ when $\epsilon = 0.3$. This implies that the attacker explores the possible actions less frequently and exploit the Q-table often to determine the best action to take in each stage, which eventually leads to a better performance.

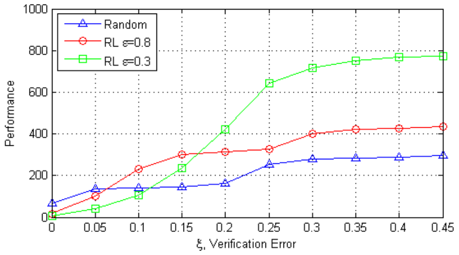

Fig. 3. Comparison of the Performance of RL based Sybil attack and Random action based Sybil attack using SRM mechanism.

The same experiment was repeated using the settings as described above but now with FC implementing our proposed DRM mechanism. Due to the effectiveness of the DRM method, the performance of the Sybil attacker remains very low. Despite Sybil attacker's use of RL, it was not capable of "fooling" the FC into concluding in its favor. Irrespective of the attack strategy adopted by the attacker, the DRM mechanism was capable of isolating the Sybil identities and assigning them low reputations. The low reputations of the identities account for the very poor performance of the attacker which is shown in Figure 6. DRM, in comparison with SRM, is a better reputation mechanism to counteract the impact of Sybil attackers on spectrum decisions.

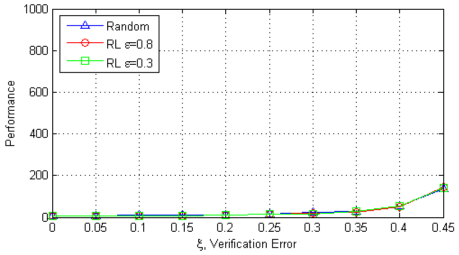

Fig. 4. Comparison of the Performance of RL based Sybil attack and Random action based Sybil attack using DRM mechanism.

B. *Comparison of Identity Sampling Strategies*

In the next set of experiments, we compare the performances of the Sybil identity sampling strategies discussed in section III C. To test their various performances we implement SRM decision making mechanism for the FC. In the first simulation setting we kept $\epsilon = 0.3$. The result of the experiment is illustrated in Figure 5. From the illustration above, it can be seen vividly that BP outperforms other sampling strategies employed by the attacker. As seen in previous experiments, an attacker experiences performance improvement as ξ increases which reflects FC's increasing inability to accurately verify its decisions. Among the identity sampling strategies, BP shows the best performance. BP strategy which involves selecting highly successfully but less frequently used identities, accounts for the improvement performance. The poor performance of the IH strategy can be attributed to the low ϵ, which actually reduces the randomness associated with attacker's exploration of possible actions with different k.

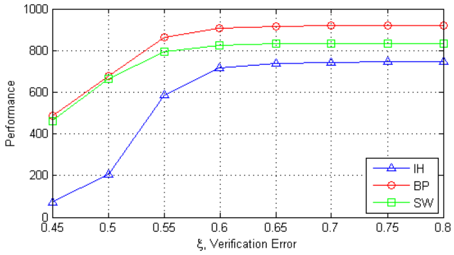

Fig. 5. Comparison of the Sybil identity sampling strategies based on their Performance with varying ξ, $\epsilon = 0.3$.

A different result was obtained in the second experiment (Figure 6), where exploration probability was increased to $\epsilon = 0.8$. BP still gives a better performance compared to other sampling strategies. However, an interesting trend is observed for IH at $\xi = 0.65$ in Figure 6, where its performance overtakes that of SW. This trend can be attributed to the increased exploration which allows the attacker to explore most of time. The increased exploratory activity of the attacker adds to the randomness that is already inherent in IH, thereby boosting its performance. We can also recall that IH is randomized BP and that explains why BP always does better than it. It can also be observed in Figure 6 that the curves produced by IH and BP are congruent. The difference between the two strategies is induced by the random selection of identities that is biased towards the best candidate identities.

C. *Comparison of Decision Making Mechanisms*

The performance of FC is of paramount importance in controlling the effectiveness of Sybil attacks. The goal, therefore, would be minimize the effectiveness of the Sybil attacks, de-

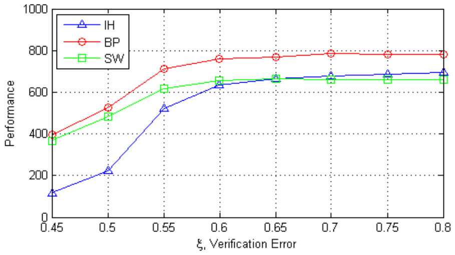

Fig. 6. Comparison of the Sybil identity sampling strategies based on their Performance with varying ξ, $\epsilon = 0.8$.

spite the strategy used by the attacker. FC assigns reputation to identities of nodes used in sensing reports and the reputation of an identity reflect the importance of its report. The reputations are used as weights to arrive at the right conclusion about the state of the spectrum.

Here we compare the performances of the static reputation mechanism (SRM) and our newly proposed dynamic reputation mechanism (DRM). The performance of a decision making mechanism is computed from the perspective of Sybil attacker's performance. Better performance of the attacker indicates a weaker mechanism to mitigate against successful attacks. Therefore we expect that the better mechanism would reduce the performance of the Sybil attacker i.e. the number of successful attacks. For experimental purpose, we kept the verification error ξ of the FC between 0 and 0.45.

The result of the first experiment with $\epsilon = 0.3$ is illustrated in Figure 7. In general, we can observe that the performance of the attacker increases as ξ increases in Figure8. But we can make clear distinction as to which reputation mechanism is best in maintaining control over Sybil identities. DRM mechanism presents a clear choice to control false reporting from the Sybil identities.

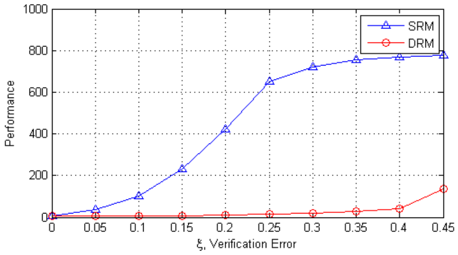

Fig. 7. Comparison of the Decision Making Mechanisms based on their performance with varying ξ, ϵ = 0.3.

The results obtained with $\epsilon = 0.8$ (see Figure 8) does not differ much from the results described above. For the SRM mechanism, we observe a decrease in the performance rate as ξ increases but DRM mechanism is more effective in keeping this rate low. We also observe that the trend for the attacker's performance using DRM mechanism is practically the same, proving the robustness of the DRM in reputation management.

The better performance of DRM mechanism can be attributed to the approach used in computing reputation for identities. The non-linear reputation function assigned separately to each identity is dynamically adjusted based on an identity's performance over the period of reporting. It is not as static and linear as implemented in the SRM mechanism. This gives the DRM mechanism an edge to control false-reporting highly effectively.

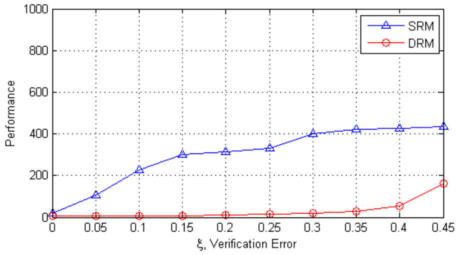

Fig. 8. Comparison of the Decision Making Methods based on their performance with varying ξ, $\epsilon = 0.8$.

VI. Conclusion

In this paper, we presented identity sampling strategies that can be employed by the Sybil attacker in choosing identities used in false reporting. The main aim of the identity sampling strategies is to minimize exposure of the identities and enhance their reputation with the FC. We demonstrated also that an attacker can improve the effectiveness of its attacks by employing RL techniques. The attacker implements Q-learning algorithm and relies on the Q-table to determine the optimal action (k identities) to perform based on its performance state. We also presented a novel dynamic reputation mechanism implemented by FC and used in keeping false reporting under control. The reputation function is robust and effective in assigning the appropriate reputation to nodes based on their performances. Our experimental findings reveal the effectiveness of the dynamic reputation mechanism making the proposed mechanism a clear defense choice against intelligent Sybil attacks.

References

[1] S. Geirhofer, L. Tong, and B.M. Sadler. Cognitive radios for dynamic spectrum access - dynamic spectrum access in the time domain: Modeling and exploiting white space. *Communications Magazine, IEEE*, 45(5):66–72, 2007.

[2] Y. Tan, K. Hong, S. Sengupta, and K.P. Subbalakshmi. Spectrum stealing via sybil attacks in dsa networks: Implementation and defense. *IEEE ICC*, 2011.

[3] Xueying Zhang and Cheng Li. Security in cognitive radio networks: a survey. *IWCMC '09: Proceedings of 2009 International Conference on Wireless Comms. and Mobile Computing*, pages 309–313, 2009.

[4] S. Anand, K. Hong, S. Sengupta, and R. Chandramouli. Is channel fragmentation/bonding in ieee 802.22 networks secure? *IEEE ICC*, 2011.

[5] Yi Tan, Shamik Sengupta, and KP Subbalakshmi. Coordinated denial-of-service attacks in ieee 802.22 networks. In *Communications (ICC), 2010 IEEE International Conference on*, pages 1–5. IEEE, 2010.

[6] James Newsome, Elaine Shi, Dawn Song, and Adrian Perrig. The sybil attack in sensor networks: analysis & defenses. In *Proceedings of the 3rd international symposium on Information processing in sensor networks*, pages 259–268. ACM, 2004.

[7] Tore Ulversoy. Software defined radio: Challenges and opportunities. *Communications Surveys & Tutorials, IEEE*, 12(4):531–550, 2010.

[8] T.C. Clancy and N. Goergen. Security in cognitive radio networks: Threats and mitigation. *Cognitive Radio Oriented Wireless Networks and Communications, 2008. CrownCom 2008. 3rd International Conference on*, pages 1–8, May 2008.

[9] Yi Tan, Kai Hong, Shamik Sengupta, and KP Subbalakshmi. Using sybil identities for primary user emulation and byzantine attacks in dsa networks. In *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, pages 1–5. IEEE, 2011.

[10] Claudia Cormio and Kaushik R Chowdhury. Common control channel design for cognitive radio wireless ad hoc networks using adaptive frequency hopping. *Ad Hoc Networks*, 8(4):430–438, 2010.