Distributed Learning for Multi-Channel Selection in Wireless Network Monitoring

Yuan Xue*, Pan Zhou†, Tao Jiang†, Shiwen Mao‡ and Xiaolei Huang*
*Department of Computer Science and Engineering, Lehigh University, Bethlehem, PA, USA
†School of Electronic Information and Communications, Huazhong University of Science & Technology, Wuhan, China
‡Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA
Email: yux715@lehigh.edu, panzhou@hust.edu.cn, Tao.Jiang@ieee.org, smao@ieee.org, huang@cse.lehigh.edu

***Abstract*—In this paper, we address an important problem in the wireless monitoring, i.e., how to choose channels with best (or worst) qualities timely and accurately. We consider both scenarios of one or more sniffers simultaneously monitoring multiple channels in the same area. Since the channel information is initially unknown to the sniffers, we shall adopt learning methods during the monitoring to predict the channel condition by a short time of observation. We formulate this problem as a novel branch of the classic multi-armed bandit (MAB) problem, named *exploration bandit* problem, to achieve a trade-off between monitoring time/resource budget and the channel selection accuracy. In the multiple sniffer cases, including partly-distributed (with limited communications) and fully-distributed (without any communications) scenarios, we take communication costs and interference costs into account, and analyze how these costs affect the accuracy of channel selection. Extensive simulations are conducted and the results show that the proposed algorithms could achieve higher channel selection accuracy than other exploration bandit approaches, hence it proves the advantages of the proposed algorithms.**

I. INTRODUCTION

To deal with the complex environment of wireless networks, dedicated devices such as *sniffers*[1] [1] are implemented to monitor the wireless channel activities (e.g., transmission rate and quality of service (QoS)). These devices collect detailed information including PHY/MAC characteristics of the wireless network, which are critical for network management and diagnosis. However, due to the limitation of hardware devices, one sniffer cannot monitor all channels in its vicinity at one time. Network monitoring is very important in many type of scenarios, for example, the IEEE 802.11 protocol divides spectrums into multiple channels, for which multiple sniffers are needed. Generally, multiple monitoring devices are deployed in the same area to cooperatively monitor channels in this area.

Unlike previous works focusing on the *channel assignment* problem [2], [3] where different sniffers are in charge of different channels, we concentrate more on the *channel selection* in wireless monitoring, i.e., how to effectively choose an optimal (or worst for network diagnose) collection of channels by each sniffer independently. Channel selection is an essential problem of wireless monitoring, for instance, choosing channels with the highest transmission rates are helpful for resource allocation and could avoid wasting time/energy on bad channels. Meanwhile, identifying channels with unusual behaviors (which may be caused by hardware failures or external attacks) in a timely and accurate fashion plays an indispensable role in some mission-critical scenarios such as voice over IP (VoIP), disaster recovery and military applications.

We investigate the problem of distributed wireless monitoring in this paper. When multiple sniffers coexist in the same area, our goal is to figure out how to coordinate these sniffers and make the optimal selection of channels without a central controller. Initially, the channel activities are unknown to the sniffers, i.e., they have to collect channel information constantly by themselves during a sequential learning process. Intuitively, if we spend more time on monitoring, we will obtain more accurate results. Unfortunately, in most cases, we do not have an infinite time for channel selection, therefore a time budget is needed. As a result, there is a trade-off between completing the monitoring quickly versus making more accurate selections. This non-trivial problem inspired us to formulate it as a *distributed exploration bandit* or *pure exploration* problem, which is a subclass of the classic multi-armed bandit (MAB) problem [4]. In contrast to standard MAB algorithms such as UCB [5], which are evaluated in terms of *regret*,[2] pure exploration methods focus on finding the arm(s) with the maximum expected rewards rather than maximizing arm rewards during the learning process.

Firstly, we study the channel selection problem with a single sniffer. Although the sniffer can only monitor one channel at a time, it can switch its target constantly. Thus, if we could predict channel condition by a short time of observation, we could use one sniffer to monitor multiple channels. We assume that there is a time budget for the sniffer, so it has to make decisions within the given time horizon. We then model this problem as a multiple identification problem with a *fixed budget* [6]. We propose an elimination-based learning algorithm, which allows the sniffer to reduce the number of channels to monitor constantly during the iterations. We show that the proposed single sniffer algorithm has a negative-

[1]Sniffer is a passive monitoring device. In this paper, for simplicity, we use the term sniffer to refer to both active and passive monitoring devices.

[2]Regret is defined as the difference between the expected rewards of the optimal strategy made by a *genie* and that of the given policy.

978-1-5090-1732-4/16/$31.00 ©2016 IEEE

exponential *error probability*[3] over time.

Next, we consider the case where multiple sniffers coexist in the same area. If they work for the same administrator, they can share some information to obtain more accurate monitoring results. Nevertheless, due to the geographical separation of the sniffers, this cooperation incurs a communication cost. For such a scenario, we propose a distributed exploration bandit algorithm with limited communications. For passive sniffers, there will not be any inferences when they observe the same channel, but they still have communication cost then they communicate with each other. As for deploying active monitoring devices to monitor the network, interferences between them may hurt their ability of getting correct monitoring results. We name this kind of interference as a *collision*. Both communication cost and collision are taken into consideration in our analysis. In another practical scenario, different sniffers might be operated by different administrators. So they cannot communicate with each other, and still have to suffer the loss of accuracy caused by collisions if they are active devices. Thus, we propose fully distributed algorithms for this scenario, in which sniffers have no need to communicate with each other.

The main contributions of this work are described below:

- We address the channel selection in wireless monitoring by formulating it as an exploration bandit problem. Both single sniffer and multiple sniffer algorithms are proposed for different scenarios.
- We investigate the distributed exploration bandit problem, mainly in the fixed budget setting. This technique could have lots of applications but hardly any work have been done in this area.
- Communication costs and collision loss are also taken into consideration in this paper. Two well-designed algorithms are proposed to mitigate the negative impact of collisions by allowing limited communications.

The rest of this paper is organized as follows. In Section II, related works on wireless monitoring and distributed learning are discussed. Section III introduces the problem formulation of distributed channel selections in wireless monitoring. Algorithms and analysis are presented in Section IV and V, respectively. Simulation results are available in Section VI. We conclude this paper in Section VII.

II. Related Works

Many previous works have been done on wireless monitoring, most of which attempted to design a complete monitoring system. In [1], for the first time, Yeo *et al.* introduced passive monitoring for the IEEE 802.11 based wireless networks with multiple sniffers. Deshpande *et al.* also studied the multi-channel wireless monitoring problem in [7], where they mainly focused on achieving higher observed frame rates with the proposed channel sampling strategies. In [2], Zheng *et al.* proposed online learning algorithms to solve the channel assignment problem of wireless monitoring in unknown environments. We also adopt online learning techniques in our work but the problem is totally different. Unlike [2], we address the channel selection problem of the wireless monitoring. We also propose a channel assignment algorithm to solve the distributed multiple channel selection problem.

Chen *et al.* in [3] also studied the channel selection problem in wireless monitoring. However, they choose channels for different sniffers to allow them to monitor multiple channels together. In their optimization problem, a central controller is required and each sniffer can always communicate with other sniffers. The central controller is not needed in our work, moreover, we consider the communication cost in distributed wireless monitoring.

In the area of online learning, multi-armed bandit (MAB) problem has drawn a lot of attention in recent decades. MAB is a classic example of tradeoff between exploration and exploitation, which aims to achieve the maximum cumulative sum of rewards in the learning process. Lai & Robbins proposed an index policy in [4] with a logarithmic regret bound and Auer *et al.* [5] introduced the well-known UCB strategy which achieves $O(\log T)$ regret uniformly over time.

Consider the large amount of channels and dynamic environment, we introduce an innovative scheme with exploration bandit techniques in this paper. Exploration bandit is a new branch of MAB and it can be mainly divided into two categories: fixed confidence setting and fixed budget setting. In the fixed budget setting , players should seek for a single best arm or a best subset of arms with a fixed time budget. In another setting of the exploration problem named fixed confidence setting, such as Even-Dar *et al.* in [8] and Kalyanakrishnan *et al.* in [9], players aim at reducing the number of samples (i.e., *simple regret* in [10]) to satisfy the specific constant of finding near-optimal arm(s). Compared with classic MAB methods, exploration bandit methods have more accurate results in the channel selection of wireless monitoring because they spend more time on bad channels. Thus we adopt exploration bandit algorithms in this paper.

In the exploration bandit problem, many works on the multiple identifications problem (EXPLORE-m in [11]) has been done such as [6] and [9]. Nevertheless, none of them has considered subset selection with different players. Different from the existing literature, we study the challenging issue of multiple identifications in the distributed setting.

There have also been considerable efforts on distributed learning techniques of MAB. In area of distributed learning, Liu and Zhao [12] introduced the time-division fair sharing (TDFS) policy for a centralized time sharing schedule for multiple users. Tekin and Liu [13] utilized the regenerative property of markov chain to solve the problem of rested and restless MAB problems with multiple players. Kalathil *et al.* in [14] proposed an algorithm based on the Bertsekas auction algorithm, which has $O(\log^2 t)$ regret bound due to the communication cost.

Distributed exploration bandit was studied by Hillel *et al.*

[3]Error probability is defined as the probability of choosing any suboptimal channels during the monitoring.

in [15]. Their work is most related to this paper. However, our work is completely independent with respect to [15]. Hillel *et al.* studied the problem of distributed exploration in the fixed confidence setting while we mainly focus on the fixed budget setting. They speed up the learning process via communications among sniffers. In contrast to their work, both algorithms with and without communications are proposed in this paper.

III. Problem Formulation

In this section, we present the problem formulation of channel selection in wireless monitoring. First, we discuss the formulation of the single sniffer and multiple sniffer problems.

A. System Model

Consider one or more sniffers monitors K wireless channels in some area. When the scale of the network K is very large, the administrator needs to focus on a subset of best or worst channels to manage the network more efficiently. Thus, the sniffers have to choose M best $(M < K)$ channels [4] out of K channels within the time budget T. According to different usage, the time budget can be different and the sniffers can finish the channel selection with or without communications with each other.

We assume that the sniffer can only monitor one channel at a time. In the single sniffer scenario, the sniffer is given a time budget T firstly, then it has to choose M channels within time T. After finishing the channel selection, the sniffer will transmit the results to the administrator of this network for further monitoring work.

In the multiple sniffers scenario, each sniffer will yield a complete outcome of M channels independently. Collisions will happen if sniffers are active monitoring devices. Assume there are n sniffers and $n \leq M$, if all the sniffers are operated by the same administrator, they can communicate during the monitoring process to avoid collision among themselves. Communication costs will hurt the reward achieved by sniffers, and further affect the accuracy of selection. Under the fully-distributed scenario, the sniffers do not communicate with each other. They will make decisions based on their past monitoring results. Also, they will choose M channels within the time budget T independently. Note that in fully distributed monitoring cases, collisions will happen.

Next, we introduce some notations and complexity measurements in this paper.

B. Notations and Complexity Measurements

Consider K channels to be monitored in a wireless network, where $\mathcal{K} = \{1, \ldots, K\}$ is the set of channels. For simplicity, we assume that each channel j's activity in wireless network follows an i.i.d. distribution with density function $f(x; \theta_j)$, while the parameter θ_j is not known *a priori*. Each time the sniffer measures the channels, it will obtain a *reward*. Reward here contains information of channel activities measured by the sniffer. When there are multiple sniffer in the system, let $\mathcal{N} = \{1, \ldots, n\}$ denote the set of sniffers. Let ϕ be the monitoring policy adopted by the sniffers. During the channel selection process, all the sniffers have the same time budget T and their clocks are synchronized. Error probability e_i^{ϕ} is the probability of choosing suboptimal channels by sniffer i under policy ϕ.

Assume that each channel j has a mean reward μ_j, which is the mean of random variable $X_j(t)$. We rank them in a descending order, i.e., $\mu_1 > \ldots > \mu_K$. We also introduce the notation of hardness H^M in this subsection for preparation of the proof. We define the gaps and the complexity measures of the distributed channel selection in wireless monitoring problem as follows.

$$\Delta_j^M = |\mu_M - \mu_j|, \tag{1}$$

$$\Delta_{\min} = \min_{1 \leq j < k \leq K} (\mu_k - \mu_j), \tag{2}$$

$$H_1^M = \sum_{j=1}^{K} \frac{1}{(\Delta_j^M)^2}, \tag{3}$$

$$H_2^M = \max_{j \in \mathcal{K}} \frac{j}{(\Delta_j^M)^2}, \tag{4}$$

where the notation $j \in \{1, \ldots, K\}$ is determined by order of $\Delta_1^M \leq \ldots \leq \Delta_K^M$. These notations decide the hardness of finding the optimal channels during the channel selection process, we will discuss more details in the following sections.

IV. Single Sniffer Monitoring

In this section, we introduce a novel single sniffer algorithm named Sequential Multiple Elimination (SME) for channel selection in wireless monitoring.

Details of the proposed algorithm are given in Algorithm 1. In SME, different from previous elimination-based algorithms, we allow sniffers to eliminate multiple channels in each round where each round divide the time budget equally. The number of eliminated channel is a quarter of last round, to let sniffers observe channels more times when it is harder to distinguish distinguish "good" channels from "bad" channels, because the reward gap between different channels become smaller as exploration time goes by. We pick a quarter because after several experiments, we found that this is a relatively good choice which is neither too big nor too small, all parameters are designed according to this principle.

First, we divide the channel selection process into l rounds and let $l = \lceil \log_4 (3(K - M) + 1) \rceil$. Thus, in round τ, the sniffer will remove $K_\tau = \lceil (3(K - M) + 1)/4^\tau \rceil$ channels with the lowest empirical rewards from the remaining active channel set $\mathcal{A}_\tau$, and put them into set $\mathcal{K}_\tau$. It follows that

$$\sum_{\tau=1}^{l} K_\tau \leq \frac{1}{3}(3(K - M) + 1)(1 - 4^{-l}) = K - M. \tag{5}$$

Let $A_\tau = |\mathcal{A}_\tau|$ and $A_\tau = \left\lceil \frac{K-M}{4^\tau} + \frac{1}{3 \cdot 4^\tau} + M - \frac{1}{3} \right\rceil$, for all $\tau \leq l$. After l rounds, the sniffer will provide the result of M chosen channels.

SME allows sniffers to reduce the number of samples constantly during the process of channel monitoring, and guarantees each channel has been sampled enough times before been dropped. To calculate the error probability of SME, we introduce a lemma first.

[4] For simplicity, we only use the optimal channels selection assumption in this paper. Note that all algorithms and analyses can be easily extended to the worst channel selection scenario.

Algorithm 1 Sequential Multiple Elimination for Single User (SME)

Input : K channels, M chosen channels, time budget T, $t_\tau = \lceil \frac{T}{lA_\tau} \rceil$, $l = \lceil \log_4 (3(K-M)+1) \rceil$, $A_\tau = |\mathcal{A}_\tau|$, $\mathcal{A}_0 = \mathcal{K}$, $K_\tau = \left\lceil \frac{3(K-M)+1}{4^\tau} \right\rceil$.
Output : M channels with highest empirical rewards.
for each $\tau = 1, 2, \ldots, l$ **do**
    Sample all channels in $\mathcal{A}_\tau$ for t_τ times;
    Rank these channels according to their empirical rewards, let $\hat{\mu}_1 > \ldots > \hat{\mu}_{A_\tau}$;
    Eliminate all the channels in $\mathcal{K}_\tau = \{j : \hat{\mu}_j \le \hat{\mu}_{A_\tau - K_\tau}\}$, $\mathcal{A}_{\tau+1} = \mathcal{A}_\tau / \mathcal{K}_\tau$.
end for
Return $\mathcal{A}_l$.

Lemma 1. *In SME, assume a channel p outside $\mathcal{M}$ is not eliminated before round τ. Then in round τ, for channel $j \in \mathcal{M}$, the probability of $\hat{\mu}_j < \hat{\mu}_p$ satisfies*

$$\mathbb{P}[\hat{\mu}_j < \hat{\mu}_p] \le \exp\left(-2\sum t_\tau (\Delta_j^M + \Delta_p^M)^2\right). \qquad (6)$$

Proof: Let $\Delta_{jp} = \Delta_j^M + \Delta_p^M$ when $j \le M$ and $p > M$. For $\alpha > 0$ and $\beta > 0$, by the Chernoff-Hoeffding inequality, we have

$$\mathbb{P}[\hat{\mu}_p > \mu_p + \alpha\Delta_{jp}] \le \exp\left(-2\sum t_\tau (\alpha\Delta_{jp})^2\right)$$
$$\mathbb{P}[\hat{\mu}_j < \mu_j - \beta\Delta_{jp}] \le \exp\left(-2\sum t_\tau (\beta\Delta_{jp})^2\right).$$

Since $\Delta_{jp} = \Delta_j^M + \Delta_p^M = \mu_j - \mu_p$, we have

$$\begin{aligned}\mathbb{P}[\hat{\mu}_j < \hat{\mu}_p] &\le \exp\left(-2\sum t_\tau ((\alpha\Delta_{jp})^2 + (\beta\Delta_{jp})^2)\right)\\ &\le \exp\left(-\sum t_\tau (\Delta_j^M + \Delta_p^M)^2\right),\end{aligned}$$

where the last inequality is due to the fact that if $\alpha + \beta \ge 1$, then $\alpha^2 + \beta^2 \ge \frac{1}{2}$. ∎

From Lemma 1 we know that under the policy of SME, when the mean reward gap between good channels and bad channels Δ_{jp} is very large, the probability for the sniffer of identifying the optimal channels becomes very high. Besides, if the sniffer spends more time on monitoring, it will achieve lower probability of choosing bad channels. Then we can derive an upper bound for the error probability of SME.

Theorem 1. *The error probability of SME is upper bounded as*

$$\begin{aligned}e \le \frac{1}{3} M \left(4K - 4M - \log_4(3(K-M)+1)\right) \times \\ \exp\left(-\frac{T}{H_2 \log_4(3(K-M)+1)}\right).\end{aligned} \qquad (7)$$

Proof: Assume that channel $j \in \mathcal{M}$ is not eliminated by SME in the first $\tau - 1$ rounds. Then in round τ, the probability for channel j of being eliminated satisfies

$$\begin{aligned}e_\tau &\le \sum\nolimits_{j\in\mathcal{M}} \sum\nolimits_{A_\tau \le p \le K} \mathbb{P}[\hat{\mu}_j < \hat{\mu}_p]\\ &\le \sum\nolimits_{j\in\mathcal{M}} \sum\nolimits_{A_\tau \le p \le K} \exp\left(-\sum t_\tau \Delta_{jp}^2\right)\\ &\le \sum\nolimits_{j\in\mathcal{M}} (A_\tau - M) \exp\left(-\sum t_\tau \Delta_{jA_\tau}^2\right)\\ &\le M(A_\tau - M) \max \exp\left(-\frac{T}{lA_\tau}(\Delta_{A_\tau}^M)^2\right)\\ &\le M(A_\tau - M) \exp\left(-\frac{T}{lH_2}\right).\end{aligned} \qquad (8)$$

Then for all rounds, we have

$$\begin{aligned}e &\le \sum\nolimits_{\tau=0}^{l} M(A_\tau - M) \exp\left(-\frac{T}{lH_2}\right)\\ &= \frac{1}{3} M(4K - 4M - l) \exp\left(-\frac{T}{lH_2}\right)\\ &\le \frac{1}{3} M \left(4K - 4M - \log_4(3(K-M)+1)\right) \times\\ &\quad \exp\left(-\frac{T}{H_2 \log_4(3(K-M)+1)}\right).\end{aligned} \qquad (9)$$

∎

Theorem 1 shows that the error probability of SME is $O(e^{-T})$ with respect to time budget T. So as the time budget grows, the error probability of SME decreases exponentially. Also, SME needs at most $O(\log K \log(K - \log K))$ times of observation with respect to number of channels K to identify the optimal channels, which is much smaller than previous algorithms such as SAR [6]. Thus, SME has better performance in the channel selection of wireless monitoring.

V. Distributed Monitoring

In this section, we examine the scenarios where multiple sniffers monitor the channels in the same area. Assume there are n sniffers monitoring K channels simultaneously, each channel's information is initially unknown to all the sniffers. Each of the sniffers will provide a complete set of M chosen channels after a period of time T. A new challenge for this multiple sniffer scenario is that when more than one active sniffers are on the same channel, the interference between them will hurt the rewards observed by these sniffers. Such collisions should be considered in the design of multiple active monitoring devices schemes.

Based on our single sniffer algorithm, we first introduce two collision-free algorithms which are for sniffers with limited communications. Then we study the fully distributed circumstance in which sniffers do not change information with each other.

A. Partly Distributed Algorithms with Limited Communications

We assume that there are n sniffers monitoring the same K channels simultaneously for the same administrator. They can communicate with each other to avoid the collision among sniffers. However, the communication cost will degrade the accuracy of the the results gained by the sniffers. So it would be costly for the sniffers to keep on exchanging information with each other. During the learning process, each sniffer has to

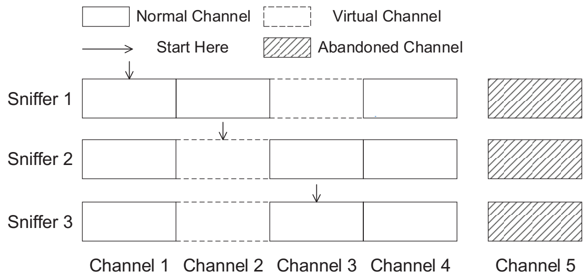

Fig. 1. Channel Selection Model of Algorithm 2.

make their own decision with limited help from other sniffers. We proposed two distributed exploration bandit algorithms in this section to solve this problem. In the proposed algorithms, communication cost for each sniffer is taken into account.

First, we propose an algorithm named Distributed Sequential Multiple Elimination with Virtual Channels (DSME-VC) in Algorithm 2. In DSME-VC, we also evenly divide the time budget T into l rounds. The elimination process is the same as the SME algorithm. Next, the sniffer will put its chosen channels into a channel set $\mathcal{V}$, which we call the *virtual channel* set. The sniffer will spend time on virtual channels but they will not collect any channel information. Meanwhile, the sniffer will broadcast its chosen channels to all other sniffers. If one channel is chosen by all the sniffers, then the sniffer will remove this channel from $\mathcal{V}$.

As illustrated in Fig. 1, with a round-robin fashion time allocation policy, when the sniffer 2 is assigned to monitor a virtual channel 2, it won't collect any information about channel 2 and just rest for time t_τ. This strategy can completely avoid the potential collisions among sniffers. For different sniffers, they may choose different channels in the same round, so we should calculate the probability for a channel of being chose by all the sniffers. Then we can compute the expectation of the remaining channels in each round.

In DSME-VC, we completely avoid the potential collisions among sniffers by introducing the virtual channels. A sniffer never really drops a channel unless it thinks the channel is selected by all sniffers simultaneously. Compared with the single sniffer algorithm, DSME-VC wastes some time on virtual channels, which is an inevitable cost for avoiding collisions. At round τ, assume the communication cost for each sniffer is c_τ. Since we divide T into l rounds, the total communication cost of l rounds is $C = \sum c_\tau$. Actually, the communication cost will hurt the results observed by the sniffers, and will affect the accuracy of monitoring results.

To calculate the communication cost of DSME-VC, we firstly define the complexity communication cost as

$$c_0 = \max_{i\in\mathcal{N}, j\in\mathcal{K}} \{c_{i,j}\}, \qquad (10)$$

where $c_{i,j}$ is the communication cost for sniffer i of exchanging information about channel j with another sniffer. c_0 here refers to the maximum communication cost for a single sniffer and single channel.

Consider n sniffers monitor multiple channels simultaneously, then sniffer i's communication cost in round τ is $c_{i,\tau} = n\sum_{j=1}^{K_\tau} c_{i,j}$. Thus, the total communication cost for sniffer i is upper bounded as

$$C_i = \sum_{\tau=1}^{l} c_{i,\tau} = n\sum_{\tau=1}^{l}\sum_{j=1}^{K_\tau} c_{i,\tau} \le n\sum_{\tau=1}^{l} K_\tau c_0 = n(K-M)c_0.$$

Algorithm 2 Distributed Sequential Multiple Elimination with Virtual Channels (DSME-VC)

Input : K channels, M chosen channels, n sniffers, time budget T, $l = \lceil \log_4(3(K-M)+1)\rceil$, $A_\tau = |\mathcal{A}_\tau|$, $\mathcal{A}_0 = \mathcal{K}$, $K_\tau = \left\lceil \frac{3(K-M)+1}{4^\tau} \right\rceil$, $k = 0$, $\mathcal{V} = \varnothing$.
Output : M channels with highest empirical rewards.
for each $\tau = 1, 2, \ldots, l$ **do**
  for each sniffer $i \in \mathcal{N}$ **do**
    while $k \le K$ **do**
      if the channel $[(i+\tau+k) \bmod K-1]$ belongs to $\mathcal{A}_\tau$ **then**
        Sample it for $t_\tau = \frac{T}{l|\mathcal{A}_\tau \bigcup \mathcal{V}|}$ times;
      end if
      $k := k+1$;
    end while
    Rank these channels according to their empirical rewards, let $\hat{\mu}_1 > \ldots > \hat{\mu}_{A_\tau}$;
    Choose all the channels in $\mathcal{K}_\tau = \{j : \hat{\mu}_j \le \hat{\mu}_{A_\tau - K_\tau}\}$, $\mathcal{A}_{\tau+1} = \mathcal{A}_\tau/\mathcal{K}_\tau$. Broadcast chosen channels to other sniffers;
    Eliminate channels chose by all sniffers, put others back to $\mathcal{V}$ as virtual channels.
  end for
end for

In the following, we calculate an upper bound for the probability for any channel j to be selected in round τ. Then we will prove the expected number of channels chosen by all the sniffers.

Lemma 2. *In DSME-VC, The probability for channel j outside K_τ to be chosen by one sniffer in round τ satisfies*

$$\mathbb{P}[j \in \mathcal{A}_{i,\tau}] \le K_{\tau-1}\exp(-\Delta_{\min}^2 T_{\tau-1}) - K_\tau \exp(-\Delta_{\min}^2 T_\tau). \qquad (11)$$

Proof: Based on Lemma 1 and a union bound, we have

$$\begin{aligned}\mathbb{P}[j \in \mathcal{A}_{i,\tau}] &\le \mathbb{P}[\textstyle\bigcup_{A_\tau < k \le A_{\tau-1}} (\hat{\mu}_j < \hat{\mu}_k)] \\ &\le \textstyle\sum_{k>A_\tau}^{A_{\tau-1}} \mathbb{P}[\hat{\mu}_j < \hat{\mu}_k] \\ &\le \textstyle\sum_{k>A_\tau}^{A_{\tau-1}} \exp(-(\Delta_{jk})^2 t_\tau) \\ &\le (A_{\tau-1} - A_\tau)\exp(-(\Delta_{jA_\tau})^2 t_\tau).\end{aligned}$$

For any $j < K_\tau$, we have

$$\begin{aligned}&\mathbb{P}[j \in \mathcal{A}_{i,\tau}] \\ &\le (A_{\tau-1} - A_\tau)\exp(-(\Delta_{jA_\tau})^2 t_\tau) \\ &\le (A_{\tau-1} - A_\tau)\exp(-(\Delta_{A_{\tau-1}} - \Delta_{A_\tau})^2 t_\tau) \\ &\le (A_{\tau-1} - A_\tau)\exp(-\Delta_{\min}^2 t_\tau) \\ &\le A_{\tau-1}\exp(-\Delta_{\min}^2 t_{\tau-1}) - A_\tau \exp(-\Delta_{\min}^2 t_\tau). \qquad (12)\end{aligned}$$

■

With the conclusion of Lemma 2, we can now prove the expected number of channels chosen by all the sniffers.

Theorem 2. *In DSME-VC, the expectation of number of channels chosen by all sniffers in the round τ satisfies*

$$\mathbb{E}[N] \geq (A_{\tau-1} - A_\tau)\left(\frac{1}{A_{\tau-1} - A_\tau} - A_\tau \exp(-\Delta_{\min}^2 \frac{T}{Kl})\right)^n \triangleq N_\tau. \tag{13}$$

Proof: For a channel inside $\mathcal{K}_\tau$, the probability of being chosen by one sniffer is at least $1 - \sum_{j=1}^{A_\tau} \mathbb{P}[j \in \mathcal{A}_{i,\tau}]$. Then we have

$$\begin{aligned}\mathbb{E}[N] &\geq (A_{\tau-1} - A_\tau)\left(\frac{1 - A_\tau \mathbb{P}[j \in \mathcal{A}_{i,\tau}]}{A_{\tau-1} - A_\tau}\right)^n \\ &\geq (A_{\tau-1} - A_\tau)\left(\frac{1}{A_{\tau-1} - A_\tau} - A_\tau \exp(-\Delta_{\min}^2 t_\tau)\right)^n, \end{aligned} \tag{14}$$

where (14) follows from the last inequality in (12). Since $t_\tau = T/l(K - \sum_{r=1}^{\tau} N_r) \leq T/Kl$, we have Theorem 2. ■

In Theorem 2, the probability for all the sniffers choose the same channel grows exponentially as the increase of the number of sniffers n. This fact guarantees that when there are many sniffers in the same area, DSME-VC will not have too many virtual channels. Next, we derive an upper bound on each sniffer's error probability for Algorithm 2.

Theorem 3. *The error probability of DSME-VC for each sniffer satisfies*

$$e \leq \sum_{\tau=1}^{l} M(A_\tau - M) \exp\left(-\frac{TM}{lH_2(K - N_1' + N_\tau')}\right), \tag{15}$$

where N_τ' *is* $A_\tau\left(1/(A_{\tau-1} - A_\tau) - A_\tau \exp(-\Delta_{\min}^2 T/Kl)\right)^n$.

Proof: First, we will modify the time allocation policy with respect to Algorithm 1.

$$T_\tau = \sum_{r=1}^{\tau} t_r \leq \sum_{r=1}^{\tau} \frac{T}{l\left(K - \sum_{r \leq \tau} N_r\right)}, \tag{16}$$

from Theorem 2 we have that

$$\sum_{r \leq \tau} N_r \geq K\left(\frac{4}{3(K-M)+1} - A_1 \exp(-\Delta_{\min}^2 \frac{T}{Kl})\right)^n - A_\tau\left(\frac{1}{A_{\tau-1} - A_\tau} - A_\tau \exp(-\Delta_{\min}^2 \frac{T}{Kl})\right)^n.$$

With (8) in Theorem 1, the error probability in round τ is upper bounded as

$$\begin{aligned} e_\tau &\leq M(A_\tau - M) \exp\left(-\frac{TA_\tau}{lH_2(K - \sum_{r \leq \tau} N_r)}\right) \\ &\leq M(A_\tau - M) \exp\left(-\frac{TM}{lH_2(K - (N_1' - N_\tau')}\right). \end{aligned} \tag{17}$$

Then, the total error probability of Algorithm 2 is

$$\begin{aligned} e &\leq \sum\nolimits_{\tau=1}^{l} e_\tau \\ &\leq \sum\nolimits_{\tau=1}^{l} M(A_\tau - M) \exp\left(-\frac{TM}{lH_2(K - N_1' + N_\tau')}\right). \end{aligned}$$

■

Compared with Theorems 1, Theorem 3 shows that the upper error probability bound of DSME-VC is bigger than SME. However, it completely circumvents the collisions among sniffers and when K is not very large, the error probability of DSME-VC is quite close to SME. So the DSME-VC algorithm is applicable in the distributed monitoring when the number of sniffers is relatively small.

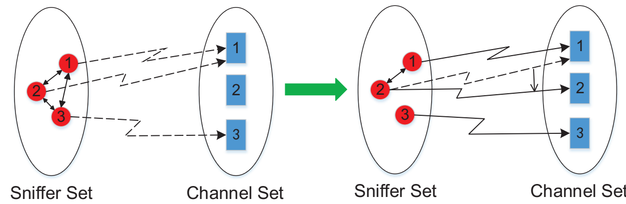

Fig. 2. The negotiation process between sniffers in Algorithm 3.

Algorithm 3 Distributed Auction-based Channel Assignment Algorithm (DACA)

1: **Input** : K channels, M chosen channels, n sniffers, time budget T, $t_\tau = \frac{T}{lK_\tau}$, $l = \lceil \log_4(3(K-M)+1) \rceil$, $A_\tau = |\mathcal{A}_\tau|$, $\mathcal{A}_0 = \mathcal{K}$.
2: **Initialization** : For a random channel j, sniffer i provides a price p_{ij} randomly, communicate p_{ij} with other sniffers.
3: **while** $\tau < l$ **do**
4:   **while** $k < A_\tau$ **do**
5:     For sniffer i, let $j = \arg\max_{j \in \mathcal{S}_{k,\tau}} \hat{\mu}_{i,j}$;
6:     **if** $i = \arg\max_{i \in \mathcal{N}} p_{ij}$ **then**
7:       Sniffer i samples channel j for t_τ times, broadcast to all other sniffers that it has finished this round, then waits for other sniffers;
8:     **else**
9:       Move to another channel $s \in \mathcal{S}_{k,\tau}$ randomly, let $p_{is} = \hat{\mu}_{i,s}$, communicate with other remaining sniffers, go back to step 6;
10:     **end if**
11:     Let $\mathcal{S}_{k,\tau} := \mathcal{S}_{k,\tau}/j, k := k+1$;
12:   **end while**
13:   Update the empirical reward $\hat{\mu}_{i,j}$ for all $j \in \mathcal{A}_\tau$. Eliminate channels in $\mathcal{K}_\tau = \{j : \hat{\mu}_j \leq \hat{\mu}_{A_\tau - K_\tau}\}$, $\mathcal{A}_{\tau+1} := \mathcal{A}_\tau/\mathcal{K}_\tau, k := 1, \tau := \tau+1, \mathcal{S}_{k,\tau} := \mathcal{A}_\tau$.
14: **end while**
15: Return $\mathcal{A}_l$.

We also propose a distributed algorithm without using virtual channels in Algorithm 3 named Distributed Auction-based Channel Assignment (DACA). Although DSME-VC solves the potential collision problem, when K becomes very large, DSME-VC may waste too much time on virtual times. So Algorithm 2 is not efficient enough when K is very large. DACA solves this problem without using the virtual channels.

Assume there is an undirected bipartite graph $\mathcal{G}(\mathcal{S},\mathcal{U},\mathcal{E})$, where $\mathcal{S}$ and $\mathcal{U}$ are the set of sniffers and channels, respectively. $\mathcal{E}$ stands for the connection between sniffers and channels. When the sniffer i eliminates channel j, the edge $E(i,j)$ will also be removed from set $\mathcal{E}$. So the sniffer will only provide a price to the channel in its active set. During the communication with others, the sniffer will decide whether to observe the channel or not. If sniffer i is not the highest bidder of channel j, it will move to another channel randomly.

For example, in Fig. 2, three sniffers are monitoring three channels. In the first communication round, both sniffer 1 and

2 provide price to channel 1 and sniffer 3 provides its price to channel 3. After communication with each other, sniffer 1 and 3 finds itself to be the highest bidder of channel 1 and 3, respectively. In the second round, sniffer 2 gives its price to channel 2, after communicate with sniffer 1, the sniffer 2 starts to collect information of channel 2. This process lasts until each of the sniffers finds a channel. After a channel is observed by channel, it be will removed from set $\mathcal{E}$ temporarily before the next round.

Since the channel selection process of DACA is the same as SME, the error probability of Algorithm 3 will also be $\left(MK - M^2 \log_4\left(12(K-M)+4\right)\right)/3 \times \exp(-T/H_2 \log_4\left(3(K-M)+1\right))$. Thus, compared with Theorem 3, the performance of DACA will be better than DSME-VC. However, to better evaluate DACA, we should take communication cost among sniffers into consideration.

Then we will derive the communication cost of DACA. First, we introduce a lemma to bound the number of communications in DACA.

Lemma 3. *Number of communications for each sniffer in DACA is at most*

$$\left(K - M + \left(M - \frac{1}{3}\right)\log_4\left(3(K-M)+1\right)\right)\frac{n^3-n}{6}.$$

Proof: We consider a worst case of Algorithm 3. When the sniffer i communicates with other sniffers in the negotiation phase of τth round, if he always fail to provide the highest price, he has to keep communicating with all remaining users, then the number of communications is at most

$$((n-1)(n-2)\cdots 1)A_\tau = \frac{n^3-n}{6}A_\tau. \tag{18}$$

Then the total number of communications is

$$\sum_{\tau=1}^{l}\frac{n^3-n}{6}A_\tau = \left(K - M + \left(M - \frac{1}{3}\right)l\right)\frac{n^3-n}{6}$$
$$\leq \left(K - M + \left(M - \frac{1}{3}\right)\log_4\left(3(K-M)+1\right)\right)\frac{n^3-n}{6}.$$

■

With the result of Lemma 3, we come to a conclusion that the communication cost of the DACA algorithm is at most $c_0\left(K - M + \left(M - \frac{1}{3}\right)\log_4\left(3(K-M)+1\right)\right)\frac{n^3-n}{6}$.

Lemma 3 shows that when the number of sniffers become very large, the DACA algorithm has much more communication cost than the DSME-VC. The numerical results of communication cost caused by DACA and DSME-VC can be found in section VI.

Since the channel selection process of DACA is the same as SME, the error probability is also the same as SME in Theorem 1. Compared with Theorem 3, DACA has lower error probability bound than DSME-VC. This is because DACA does not use virtual channels so sniffer can sample each channel for more time than DSME-VC, thus it yields more accurate results.

B. Fully Distributed Algorithms without Communications

Under the fully distributed scenario, all the sniffers cannot have any communications during the monitoring process. This maybe because they are on a military mission where communications are dangerous or they just work for different administrators so cannot exchange information with each other.

Algorithm 4 Fully Distributed Sequential Multiple Elimination (FDSME)

Input : K channels, M chosen channels, n sniffers, time budget T, $t_\tau = \lceil \frac{T}{lK_\tau} \rceil$, $l = \lceil \log_4\left(3(K-M)+1\right) \rceil$, $A_\tau = |\mathcal{A}_\tau|$, $\mathcal{A}_0 = \mathcal{K}$, $K_\tau = \left\lceil \frac{3(K-M)+1}{4^\tau} \right\rceil$.
Output : M channels with highest empirical rewards.
for each $\tau = 1, 2, \ldots, l$ **do**
  while $m < |\mathcal{A}_\tau|$ **do**
    Choose one channel in $\mathcal{A}_\tau$ randomly, sample it for t_τ times;
    Choose another channel randomly, $m := m + 1$;
  end while
  Rank these channels according to their empirical rewards, let $\hat{\mu}_1 > \ldots > \hat{\mu}_{A_\tau}$;
  Eliminate all the channels in $\mathcal{K}_\tau = \{j : \hat{\mu}_j \leq \hat{\mu}_{A_\tau - K_\tau}\}$, $\mathcal{A}_{\tau+1} = \mathcal{A}_\tau / \mathcal{K}_\tau$, $m := 1$.
end for
Return $\mathcal{A}_l$.

In this subsection, we propose two fully distributed algorithm without any communications. We first propose an algorithm named Fully Distributed Sequential Multiple Elimination (FDSME) in Algorithm 4. The idea of FDSME is intuitive: just adapt SME to the distributed setting. When n sniffers monitor K channels of some network, they just pick channels in their own active channel set following a random order. In FDSME, when multiple sniffers are active monitoring devices, collisions among them may happen. Then we prove the upper bound of collision probability given in Lemma 4:

Lemma 4. *In FDSME, the probability of at least one channel is chose by multiple sniffers simultaneously satisfies*

$$\mathbb{P}[c] \leq (\mathcal{C}_j + 1)^n - (\mathcal{C}_j)^n - n(\mathcal{C}_j)^{n-1}, \tag{19}$$

where $\mathcal{C}_j$ is $K\exp\left(-\Delta^2_{\min}t_1\right) - M\exp\left(-\Delta^2_{\min}t_l\right)$.

Proof: First, assume that there are no collisions before round τ. In the τth round, when some channel j outside $\mathcal{K}_\tau$ is chosen by more than one sniffer, a collision happens. According to Lemma 2 and a union bound, the probability for each sniffer to choose channel j outside $\mathcal{K}_\tau$ satisfies

$$\mathbb{P}[j \in \mathcal{A}_{i,\tau}] \leq A_{\tau-1}\exp(-\Delta^2_{\min}t_{\tau-1}) - A_\tau\exp(-\Delta^2_{\min}t_\tau).$$

Then we have that each sniffer's probability of choosing channel j at least once in all rounds is

$$\mathbb{P}\left[\bigcup\nolimits_{\tau \leq l}(j \in \mathcal{A}_{i,\tau})\right] \leq \sum\nolimits_{\tau<l}\mathbb{P}[j \in \mathcal{A}_{i,\tau}]$$
$$\leq K\exp(-\Delta^2_{\min}t_1) - M\exp(-\Delta^2_{\min}t_l) \triangleq \mathcal{C}_{i,j}. \tag{20}$$

Next, we calculate the collision probability for channel j in all rounds:

$$\mathbb{P}[j] \leq \mathbb{P}[\bigcup\nolimits_{k\in\mathcal{N}}(j \in \bigcap\nolimits_{i=2}^{k}\mathcal{A}_i)] \leq \sum\nolimits_{i=2}^{n} C_n^i(\mathcal{C}_{i,j})^i$$
$$= (\mathcal{C}_j + 1)^n - (\mathcal{C}_j)^n - n(\mathcal{C}_j)^{n-1}.$$

Algorithm 5 Fully Distributed Sequential Multiple Elimination with Virtual Channels (FDSME-VC)

Input : K channels, M chosen channels, $i \in \mathcal{N}$, time budget T, $t_\tau = \lceil \frac{3T/K+1}{4^{\tau-1}} \rceil$, $l = \lceil \log_4(\frac{3T}{K}+1) \rceil$, $k=1$, $\mathcal{A}_1 = \mathcal{K}$, $A_\tau = |\mathcal{A}_\tau|$, $K_\tau = \lceil \frac{K-M}{l} \rceil$, $\mathcal{V} = \varnothing$..
Output : M channels with highest empirical rewards.
for each $\tau = 1, 2, \ldots, l$ **do**
  while $k \le K$ **do**
    Set $j = [(i + \tau + k) \bmod K - 2]$
    if $j \in \mathcal{A}_\tau$ **then**
      Sniffer i samples channel j for t_τ times.
    else
      Sniffer rests for t_τ time.
    end if
  end while
  Rank these channels according to their empirical rewards, let $\hat{\mu}_1 > \ldots > \hat{\mu}_{A_\tau}$.
  Eliminate all the channels in $\mathcal{K}_\tau = \{j : \hat{\mu}_j \le \hat{\mu}_{A_\tau - K_\tau}\}$, $\mathcal{A}_{\tau+1} = \mathcal{A}_\tau / \mathcal{K}_\tau$, $\mathcal{V} := \mathcal{V} \cup \mathcal{K}_\tau$.
end for
Return $\mathcal{A}_l$.

■

Now we can calculate the error probability of FDSME.

Theorem 4. *In FDSME, the error probability for each user under fully distributed scenario satisfies*

$$e \le (l-1)M(K-M)\times \exp\left(-\frac{T}{lH_2((\mathcal{C}_j+1)^n - (\mathcal{C}_j)^n - n(\mathcal{C}_j)^{n-1})}\right). \quad (21)$$

Proof: When a collision happens, we assume that it will hurt the channel rewards observed by the sniffer. According to Lemma 4, the number of conflicting channel in round τ is at most $K_\tau(K_{\tau-1}\exp(-\Delta_{\min}^2 T_{\tau-1}) - K_\tau \exp(-\Delta_{\min}^2 T_\tau))$, so there are at most $K(\mathcal{C}_j+1)^n - (\mathcal{C}_j)^n - n(\mathcal{C}_j)^{n-1}$ conflicting channels in the whole process. Then we modify Theorem 1 to get that

$$e_\tau \le M(A_\tau - M)\exp\left(-\frac{T}{lH_2\mathcal{C}_\tau}\right) \le M(A_\tau - M)\times \exp\left(-\frac{T}{lH_2((\mathcal{C}_j+1)^n - (\mathcal{C}_j)^n - n(\mathcal{C}_j)^{n-1})}\right).$$

Thus, the total error probability is

$$\begin{aligned} e &\le \sum_{\tau=0}^{l} e_\tau \le \sum_{\tau=0}^{l} M(A_\tau - M)\times \exp\left(-\frac{T}{lH_2((\mathcal{C}_j+1)^n - (\mathcal{C}_j)^n - n(\mathcal{C}_j)^{n-1})}\right) \\ &\le \frac{1}{3}M\left(4K - 4M - \log_4(3(K-M)+1)\right)\times \exp\left(-\frac{T}{lH_2((\mathcal{C}_j+1)^n - (\mathcal{C}_j)^n - n(\mathcal{C}_j)^{n-1})}\right). \end{aligned} \quad (22)$$

■

When K and n is small, the FDSME algorithm has almost the same performance as SME. However, as the number of channels K becomes considerable, the collision probability of FDSME cannot be ignored anymore. Still, the error probability of FDSME is acceptable even when K is very large.

We also proposed another fully distributed algorithm with virtual channels in Algorithm 5 named Fully Distributed Sequential Multiple Elimination with Virtual Channels (FDSME-VC). The channel choosing process is similar to DSME-VC but the sniffer won't broadcast its monitoring results after each round. After a channel is chosen by the sniffer, it will be labeled as a virtual channel.

Time budget for each round τ is not the same. Assume the sniffer sample all channels for t_τ times in the τth round, $t_1 = \lceil 3T/K + 1 \rceil$ and $t_\tau = \lceil t_1/4^{\tau-1} \rceil$. Then we have that $K\sum_{\tau=1}^{l} t_\tau \le T$, so we can deduce that $l = \lceil \log_4(3T/K + 1) \rceil$. In τth round, when the sniffer is assigned to monitor a virtual channel, it will rest for t_τ time and does nothing. After l rounds, the sniffer yields a complete result of M chosen channels.

Theorem 5. *The error probability of FDSME-VC for each sniffer is upper bounded as*

$$e \le \frac{(\log_4(3T/K+1) - 1)(K-M)M}{2}\exp\left(-\frac{4T}{K}\Delta_{\max}^2\right), \quad (23)$$

where $\Delta_{\max} = \max_{j \in \mathcal{K}}\{\Delta_j^M\}$.

Proof: The proof of Theorem 5 is similar to Theorem 1. We omit the proof due to the space limitation. ■

Compared with FDSME, the upper error probability bound of FDSME-VC is higher when K is not considerable. However, FDSME-VC will waste much time on virtual channels when K is very large, so it has worse performance in the large scale network deployments.

VI. SIMULATION RESULTS

In this section, we show the simulation results of the proposed algorithms in wireless monitoring. We compare our single sniffer algorithm with the SAR algorithm in [6], then we illustrate the performance of the proposed distributed algorithms. We consider a few different setups where number of channels and users varies from one to another. With out loss of generality, we assume that each channel's reward is associated with an i.i.d. Bernoulli distribution. When $K = 49, M = 10$ and $K = 99, M = 15$, the parameters of each distribution are $\Theta = (0.02, 0.04, 0.06, \ldots, 0.98)$ and $\Theta = (0.01, 0.02, 0.03, \ldots, 0.99)$, respectively.

The regret here is defined as the difference between true means of optimal M channels and that of channel chosen by the algorithms. In Fig. 3, SME has much smaller regret when compared with SAR, especially when K is large. When $K = 99$, compared with SAR, SME improves the accuracy and efficiency for more than one hundred times. Even when $K = 49$, the proposed SME algorithm is ten times better than SAR.

In Fig. 4, we can see that the DACA algorithm has lower regret than DSME-VC in both scenarios. When $K = 99$, compared with DSME-VC, DACA improves the accuracy of channel selection for more than 30%. Then we compare the communication cost for two algorithms in Fig. 5.

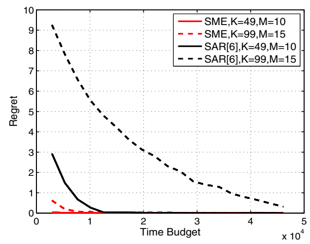

Fig. 3. Regret performances comparison for single user.

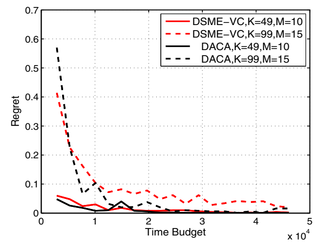

Fig. 4. Regret performances comparison for multiple users with communications.

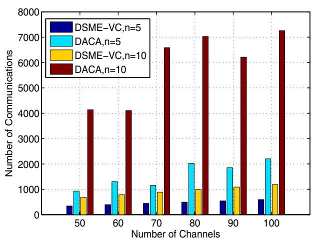

Fig. 5. Expected communication cost for each sniffer.

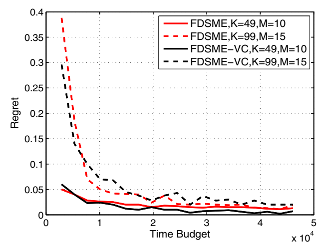

Fig. 6. Regret performances comparison for multiple users without communications.

In Fig. 5, when the number of sniffers is 5, the different between two algorithms is not very large, which indicates that the communication cost of DACA is acceptable for relatively small number of sniffers. However, when the n equals to 10, compared with DSME-VC, the communication frequency of DACA becomes very large, which also shows the advantage of DSME-VC for large scale monitoring system.

In Fig. 6, we assume there are 10 sniffers monitor K channels simultaneously. When K is relatively small, the FDSME-VC algorithm outperforms the FDSME algorithm slightly. However, when K becomes larger, FDSME wins. When $K = 99$, the accuracy of FDSME is about 10% higher than FDSME-VC. This is because FDSME-VC wastes too much time on virtual channels when K is very large.

The numerical results suggests that as the time budget grows, error probabilities for all algorithms decrease exponentially, which is completely in conformity with our theoretical analysis. More importantly, each algorithm has its advantages in specific scenarios. In summary, the simulation results prove the advantages of the proposed algorithms in the channel selection of wireless network monitoring.

VII. Conclusion

By modelling channel selection problem in wireless monitoring as an exploration bandit problem, we studied both single sniffer and multiple sniffer monitoring scenarios. As illustrated in Table I, a few single or distributed exploration bandit algorithms are proposed in this paper for different practical scenarios. Simulations are conducted and the results illustrated the performance of the proposed algorithms in the multiple channel selection. Both theoretical analysis and simulation results show that the well-designed algorithms proposed in this paper have excellent performances in different scenarios of channel selection in wireless monitoring. As far as we know, no similar work has been done before, so the proposed algorithms could have great potential applications in practice.

TABLE I
Summary of the Proposed Algorithms

Algorithm	*Distributed?*	*Communications?*	*Collisions?*
SME	×	×	×
DSME-VC	✓	✓	×
DACA	✓	✓	×
FDSME	✓	×	✓
FDSME-VC	✓	×	×

Acknowledgments

This research was supported by the National Science Foundation of China under Grant 61401169 and NSF ECCS-1509212.

References

[1] J. Yeo, M. Youssef, and A. Agrawala, "A framework for wireless LAN monitoring and its applications," in *Proc. 3rd ACM Workshop on Wireless security*, Philadelphia, PA, Oct. 2004, pp. 70–79.

[2] R. Zheng, T. Le, and Z. Han, "Approximate online learning for passive monitoring of multi-channel wireless networks," in *Proc. IEEE INFOCOM 2013*. Turin, Italey: IEEE, Apr. 2013, pp. 3111–3119.

[3] X. Chen, Y.-A. Kim, B. Wang, Y. Song, H. Dinh, and G. Chen, "Sniffer channel selection for monitoring wireless lans," *Elsevier Computer Communications*, vol. 35, no. 16, pp. 1994–2003, 2012.

[4] T. L. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules," *Elsevier Advances in Applied Mathematics*, vol. 6, no. 1, pp. 4–22, 1985.

[5] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Machine Learning*, vol. 47, no. 2-3, pp. 235–256, May 2002.

[6] S. Bubeck, T. Wang, and N. Viswanathan, "Multiple identifications in multi-armed bandits," in *Proc. 30th International Conference on Machine Learning (ICML-13)*, Atlanta, GA, June 2013, pp. 258–265.

[7] U. Deshpande, T. Henderson, and D. Kotz, "Channel sampling strategies for monitoring wireless networks," in *Proc. WiOPT 2006*. Boston, MA: IEEE, Apr. 2006, pp. 1–7.

[8] E. Even-Dar, S. Mannor, and Y. Mansour, "Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems," *Journal of Machine Learning Research*, vol. 7, pp. 1079–1105, Dec. 2006.

[9] S. Kalyanakrishnan, A. Tewari, P. Auer, and P. Stone, "Pac subset selection in stochastic multi-armed bandits," in *Proc. 29th International Conference on Machine Learning (ICML-12)*, Edinburgh, Scotland, June 2012, pp. 655–662.

[10] S. Bubeck, R. Munos, and G. Stoltz, "Pure exploration in multi-armed bandits problems," in *Algorithmic Learning Theory*. Springer, 2009, pp. 23–37.

[11] S. Kalyanakrishnan and P. Stone, "Efficient selection of multiple bandit arms: Theory and practice," in *Proc. 27th International Conference on Machine Learning (ICML-10)*, Haifa, Israel, June 2010, pp. 511–518.

[12] K. Liu and Q. Zhao, "Distributed learning in multi-armed bandit with multiple players," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5667–5681, 2010.

[13] C. Tekin and M. Liu, "Online learning of rested and restless bandits," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5588–5611, 2012.

[14] D. Kalathil, N. Nayyar, and R. Jain, "Decentralized learning for multiplayer multiarmed bandits," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2331–2345, Apr. 2014.

[15] E. Hillel, Z. S. Karnin, T. Koren, R. Lempel, and O. Somekh, "Distributed exploration in multi-armed bandits," in *Advances in Neural Information Processing Systems*, 2013, pp. 854–862.