Efficient SINR Estimating with Accuracy Control in Large Scale Cognitive Radio Networks

Yanchao Zhao†, Jie Wu*, and Sanglu Lu†
†State Key Lab of Novel Software,
Department of Computer Science and Technology, Nanjing University, China
Email: yczhao.nju@gmail.com, sanglu@nju.edu.cn
*Department of Computer and Information Sciences, Temple University, Philadelphia, PA 19122
Email: jiewu@temple.edu

***Abstract*—Recently, the SINR-model has been widely utilized in link scheduling, spectrum allocation and other applications. The SINR model requires the receiving power information of all potential link peers, which is usually assumed to be known as priori or following a uniform propagation model. We have performed experiments to illustrate how real power data could improve the performance of the SINR-based applications with considerable margin. Thus, obtaining the real power data through measurements is promising. However, this method faces many challenges. We propose a pathloss model based solution, including a representative link selection method to cut down the measurement pairs; accuracy control to determine the sample size; and a measurement distribution method to shorten the measurement duration. Our experiments show that our solution significantly improves the SINR-based scheduling's performance.**

***Index Terms*—SINR estimation, throughput optimization, pathloss model, centralized algorithm, spectrum assignment.**

I. Introduction

Recent years have witnessed the booming of wireless networking, and the emergence of new wireless network paradigms like *Cognitive Radio Networks* (CRNs) and *Wireless Mesh Networks* (WMNs). These networks provide the user with higher access speeds and better wireless resource utilization.

WMNs are considered to be a promising solution for the support of low-cost broadband Internet access for large areas [1]. A WMN consists of mesh clients and mesh routers. Mesh routers form the backbone of the network to provide network access for mesh clients. Some of them are gateways that are directly connected with the Internet via high-capacity cables. To efficiently deliver a high volume of traffic between the Internet and those non-gateway mesh routers over wireless channels, the limited bandwidth needs to be fairly allocated to them.

Cognitive radios are desirable for a WMN in which a large volume of traffic is expected to be delivered since they are able to utilize available spectrums more efficiently, thus significantly improving the network capacity [2]. However, they introduce additional complexities to bandwidth allocation. In a traditional 802.11-based WMN, a set of homogeneous channels are always available to every mesh router. Meanwhile, in a WMN with cognitive radios, each node can access a large number of spectrum bands (channels), which may spread over a wide range of frequencies. Different channels can support quite different transmission ranges and data rates, which will have a significant impact on route and channel selections.

Although CR technology provides a wide spectrum, the spectrum utilization could be further enhanced by spatial and temporal reuse. Thus, spectrum allocation that is subject to interference has been constantly drawing the attention of researchers. Recently, due to the accurate modeling of accumulative effect of interference, the SINR interference model, rather than the protocol model, is preferred. The accumulative effect implies that a transmission could fail due to a far away simultaneous transmission. With the SINR of each of the sending and receiving peers known for all the possible concurrency scenario, perfect scheduling and spectrum allocation will be available [3]–[6].

However, most of the SINR-based link scheduling and spectrum allocation algorithm assumes that either the receiving power is known as a priori, or the signal propagation follows the distance exponential attenuation model, based on which, receiving power could be directly derived by the formula containing the sending power data and the distance between nodes. Though it is typically assumed in analysis and design problems that the *Path Loss Exponent* (PLE) is known as a priori, it is often not the case. In this paper, we argue that this assumption and model could highly degrade the scheduling and assignment algorithms's performance.

To illustrate our motivation, we perform an experiment on real data. Throughput optimization algorithms based on the SINR model are conducted using both propagation model and real power information. A CDF result is shown in Fig. 1, which illustrates that there is a big margin between the line of the propagation model and the line with the power information. The results of average throughput also give out a 24% performance gap. Based on above results, we believe that there is a great opportunity to further enhance the network performance by accurately estimating the SINR value of each node under a certain transmitting mode.

This degradation is mainly caused by the power propagation model which is only applicable in long-distance signal propagation. Also, the PLE will be different when the environment

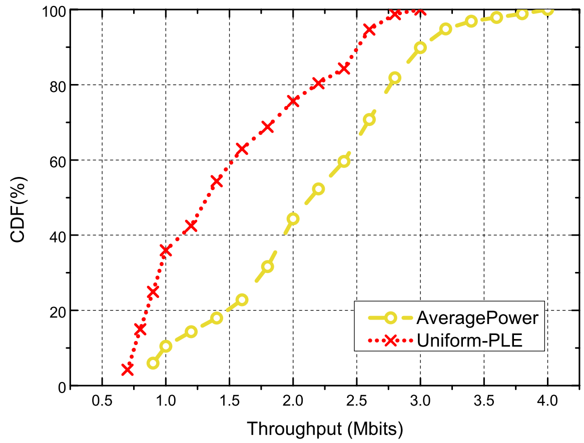

Fig. 1. The effect of applying the real power information versus using the propagation model based on distance and a uniform PLE

changes. As a result, we intuitively use measurements to get better results. However, the measurement method will face the following challenges:

- Due to shadow fading, signal power measurements will be diversified even between different measure slots of a single link.
- Just like the communication in wireless networks, we also have to make sure each measurement would not conflict with others.
- In CRN, the available spectrum spans over a large range, which introduces the spectrum diversity. In our problem, spectrum diversity means that the same signal in different bands propagate differently. Thus, our method must take this feature into consideration. But, it is also too costly to measure for every peer in every available band.

Regarding these challenges, a highly efficient and reliable SINR estimating method with accuracy control and cost control is required. Our contributions could be summarized as:

- A pathloss model-based parameter estimation method, including representative link selection and crossband power estimation.
- We introduced an accuracy control method which explores and guides the tradeoff between accuracy and measurement cost.
- We also defined the measurement optimization problem considering different targets. These problems are modeled as integer optimization problems, thus could be solved approximately.

The rest of this paper is organized as follows: in Section II, we introduce related papers. The system model and define our problems are presented in Section III. The signal propagation model based estimation and measurement methods are proposed in IV. The accuracy control method is introduced in Section V. The measurement distribution problems and its solution are proposed in Section VI. Section VII is the statement of our experiment methods and results. Finally, we conclude this paper in Section VIII.

II. Related Work

A. SINR-based Network Optimization

The SINR model is widely regarded as a better model for interference characterization. Although such a model is preferred, there are many difficulties in carrying out analyses with this model due to the computational complexity SINR involves. As a result, many previous efforts were done on single-hop networks, e.g., [3], [4]. For multi-hop networks, some efforts study cross-layer problems involving two layers instead of three layers (physical, link and network). For example, in [5], Bhatia and Kodialam optimized power control and routing, but assumed some frequency hopping mechanism is in place for scheduling, which helps simplify joint consideration of scheduling. For cross-layer optimization in the SINR model, nearly all existing efforts (e.g., [7]) followed a layer-decoupled approach to simplify analysis. Under such an approach, the solution is obtained by determining an algorithm/mechanism for one layer at a time and then piecing up them together instead of solving a joint optimization problem. Due to decoupling in the solution procedure, these approaches are heuristic at best and cannot offer any performance guarantee. Different from these works, we provide a performance-guaranteed, centralized solution along with a fast localized solution.

When compared to the traditional wireless networks, channel assignments in CRNs have to deal with the different scope of spectrum availability. Thus, various distributed approximations were proposed, which are based on observing local interference patterns [8], local bargaining [9] or on coordinations between CR nodes that aim at maximizing some system utility [10], [11].

B. Signal Fading and Pathloss Models

In most of the prior literature on PLE estimation algorithms, authors have assumed a simplified channel model consisting only of a large-scale path loss component and a shadowing counterpart; therefore, their methods have focused mainly on RSS (receive signal strength)-based localization techniques. We are, however, not aware of any related work that has considered fading, and most importantly, interference in the system model. Estimation based on a known internode distance probability distribution is discussed in [12]. The authors assume that the distance distribution between two neighboring nodes, i and j, is known or can be determined easily.

In [13], the authors consider a network where the path loss between a few low-cost sensors is measured and stored for future use. They propose an algorithm that employs interpolation techniques to estimate the path loss between a sensor and any arbitrary point in the network. In [14], a PLE estimator, based on the method of least squares, is discussed and used in the design of an efficient handover algorithm. However, as described earlier, the situation is completely different when interference and fading are considered and we cannot use these purely RSS-based estimators. Regarding the accuracy control method, in [15], the author provide a framework to control the accuracy for the measure of SINR-PRR relation.

Our work draw the lessons from these researches and work on the topic of the SINR estimation to assist the scheduling and spectrum allocation. To the best of our knowledge, this paper is the first try to solve this problem.

III. PRELIMINARIES

We first describe the system model and the assumptions in this section. Then, the problem we study in this paper is defined.

A. System Model

1) Network Model: We consider a wireless mesh network with a set of CR mesh routers $\mathcal{N}$. Each node $i \in \mathcal{N}$, senses its environment and finds a set of available spectrum bands $\mathcal{M}_i$, for the given time instance (i.e., those bands that are currently not being used by primary users), which may not be the same as the available spectrum bands at other nodes. Without the loss of generality, we assume that the bandwidth of each spectrum band (channel) is W. Denote $\mathcal{M}$ as the union of all spectrum bands among all of the nodes in the network, i.e., $\mathcal{M} = \bigcup_{i\in\mathcal{N}} \mathcal{M}_i$, and each band is identically denoted as m. We also denote $\mathcal{M}_{ij} = \mathcal{M}_i \bigcap \mathcal{M}_j$, which is the set of common bands between nodes i and j.

2) SINR Model: Let P denote the transmitting power and d_{ij} be the distance between the transmitter i and j. Let p_{ij}^m denote the received power level at node j from node i in band m; this quantity is a random variable due to wireless fading. An implicit practical assumption throughout is that $p_i > R_{th}$, where R_{th} is a positive constant for the minimum power level required to receive a signal.

In SINR model, concurrent transmissions are allowed and interference (due to transmissions by non-intended transmitters) is treated as noise. A transmission is successful if and only if the SINR at the receiver is greater than or equal to a threshold.

$$s_{ij}^m = \frac{p_{ij}^m}{N_0 + \sum_{k\in\mathcal{N},k\neq i} p_{kj}^m}. \qquad (1)$$

Here, N_0 is the ambient Gaussian noise density.

3) Signal Propagation Models: To better understand the signal propagation characteristics, we introduce the most crucial propagation model here.

There are many commonly used signal propagation models, namely, the path loss law of "large-scale effects", shadow fading of log-normal distribution and Rayleigh fading which models the multipath effect. Since we are concerned with an average signal power level, we do not consider short-term fading, which is used primarily for analyses at the data-bit level (e.g. Rayleigh and log-normal fading). Thus, to assist our model-driven efficient measurement, we choose path loss model.

Generally, the pathloss model could be depicted as signal strength attenuating with distance d as d^γ, which is:

$$S \propto (\frac{d}{d_0})^{-\gamma}, \qquad (2)$$

TABLE I
SIMULATION SETTINGS

Environment	Path Loss Exponent
Free Space	2
In building Line of Sight	1.6 to 1.8
Urban Cellular Radio	2.7 to 3.5
Obstructed in Building	4 to 6

where S is the signal strength; d_0 is the reference signal strength; and γ is the PLE. The empirical value of the PLE in different environments is shown in Table III-A3.

Let p_0 be the received power level at distance d_0. The received power at node j at distance d_{ij} from the transmitter i is then given as:

$$p_{ij} = p_0 + 10\gamma log(\frac{d_{ij}}{d_0}) \qquad (3)$$

A main feature of the large-scale path loss law is its omit of power variance. This model is also featured as deterministic model.

The reason why we apply the pathloss model to assist the measurement is mainly due to following reasons:

- First, the pathloss model is useful. It reveals the implicit correlation between different link pairs. As we know that, in the pathloss model, PLE is certain for one plain environment. We could try to find the links that share the same or a close PLE.
- Second, the pathloss model is simple. In its most simple form, it is a linear function (as in equation 3). Thus, it is quit convenient for us to utilize this model and perform accuracy control on it.

To some extent, we do not treat the path loss model in our work as its original means, because we assume the PLE is different from link to link.

B. Assumptions

To make our problem tractable and ease our solution design and analysis, we make the following assumptions:

1) We assume that the system operates synchronously in a time-slotted mode, so that each received signal and received packets will be fully overlapped. This will ease the analysis hereafter.
2) We also assume the power additive feature. This means when node i receives the signal from node j in power of p_{ji}, while the signal from k is p_{ki}. Then when node j and k send packets simultaneously to i, the power level would be $p_{ji} + p_{ki}$.
3) We assume that the position of each node is known, so that the path loss model relying on the distance the signal traveled could be utilized.
4) The interference and noise from other networks are assumed to be known, thus each node could distinguish the measurement signal from others.

C. Problem Definition and Basic Ideas

Basically, we want to get p_{ij}^m, in each node pair ij, in each of its available band m. There are two fundamentally opposite solutions to this problem. Firstly, as was applied in most of the previous algorithms, we utilize the multiple propagation models and measurements to characterize the spectrum.

On the other hand, a solution could be sequentially broadcast in every node, one band per time-slot, so that each node will measure the received power and record it.

However, the former solution suffered from inaccurate estimation, while the latter solution is too costly, especially when applied in CRNs where the available spectrum is much wider. Also, this solution has to be worked out in a fully coordinated network.

Our basic ideas are:

- To make the estimation more accurate, we try to estimate the signal propagation characteristic in every possible transmission link. The path loss model will be utilized to assist with this estimation. Because, this model reveals the signal propagation correlation of each link. Several "representatives" could be selected to perform measurement, thus cutting the measurement cost.
- The measurement overhead could be further degraded by selecting a representative band in each link. The propagation characteristic in other bands could be derived from the measured band.
- Careful scheduling of the measurement could fully utilize the spatial and temporal diversity and save the measurement cost in terms of time.

After combing the above three ideas, we thus develop an efficient SINR estimation solution in a large scale CRN. The details are presented in the following sections.

IV. Model-based Estimation

In this section, we introduce how to utilize the pathloss model to assist with our estimation and to depress measurement overhead. After this, we present how to use the cross band model to select representative band.

A. Pathloss Model-based Estimation

We use the Maximum Likelihood method to utilize the measured data and derive the corresponding γ, which is the PLE estimation problem.

The PLE estimation problem is essentially tackled by equating the empirically (observed) measured values of the aforementioned network characteristics to the theoretically established ones to obtain $\hat{\gamma}$. In each time slot, nodes either transmit or listen to recorded measurements. Upon obtaining the required measurement values over several time slots, the estimation process can be performed at each node in the network in a distributed fashion.

- Record the values of the received powers R_1, R_2, $R_3 .....R_N$.Then, take the signal power $S_i, 1 \leq i \leq N$, to be N independent realizations of an exponential random variable with unit mean.
- Use regression with the MSE rate to estimate γ.

B. Selection of representative Links

The objective of choosing the representative links is such that we can achieve the desired accuracy while a performing measurement in the minimum number of links. The modeling accuracy can be characterized by MSE in linear regression. Thus the goal is to choose the minimum number of representative links to satisfy the maximum tolerable MSE. Given that certain links may share the same or close PLE (γ), we perform the clustering method to identify the representative links. To identify the representative links, we need to

1) First, take a small number of measurements for each link;
2) Second, use the method introduced in last sub-section to estimate the γ for each link;
3) Then, use the predetermined k to divide the links into k groups with close γ;
4) Use the k-means clustering algorithm [16] to further divide the group;
5) Select the link with closest γ to the group mean γ;

Note that, here we use the k-means algorithm to perform clustering, where k is the group number. Thus we can easily control the number of representative links.

Once the representative links have been identified, they will be used to perform further accurate measurement until the accuracy of the model exceeds the required bound due to the temporal variation.

C. Crossband Signal Propagation Estimation

To make use of the measurements in different frequencies, we study how to extend the result in one channel to other channels.

In free space, the received power at a distance d from a point source could be modeled [17]:

$$PL(d) = 10log_{10}f_c^2 + 10\eta log_{10}d + L_f(n) - 28 \qquad (4)$$

where f_c is the carrier frequency, η is the path loss coefficient, $L_f(n)$ is the additional loss due to the number of floors n, between the transmitter and receiver. Thus, there is an f_c^2 dependence of the pathloss on the frequency. Some recent studies have also shown that the frequency dependence is larger than a square law in some cases, however for this paper we will assume an f_c^2 dependence which we believe to be largely true.

Suppose $P_r(f1)$ is the received power at n_2 when the transmission happens using carrier frequency f_1, and similarly $P_r(f2)$ is defined. Using Equation 2, it can be easily shown that:

$$P_r(f_1) - P_r(f_2) = 20log(\frac{f_2}{f_1}). \qquad (5)$$

Without loss of generality, assume $f1 < f2$. The equation shows that if either $Pr(f1)$ or $Pr(f2)$ is known, the other can be inferred. From this equation we can see that a measurement on one band could derive the power propagation characteristic of another band of the same link.

V. ACCURACY CONTROL

For measurements in each link peer, we further develop an accuracy control method to explore the tradeoff between measurement accuracy and overhead. In this section, we present our measurement control method.

Our accuracy control method is consisting of the control of two part: first is the control of number of representative links and the number of band pick to take measurement in each representative link. Regarding these two parameters, we will show a guidance with the experimental result in section VII; the second part is the control of the sample size in each selected band of representative links. As we have mentioned, the receiving power could be diverse due to shadow fading, multiple samples should be taken to ensure more accurate regression. In this section, we mainly discuss the accuracy control with the sample size.

Given a selected set of representative links, the overall accuracy can be ensured by controlling the measurement accuracy corresponding to the representative links and by controlling the sample size. For each representative links l_{ij} , we first compute the minimum number of signal power samples, denoted by $m_{l_{ij}}$, that we need to ensure the measurement accuracy; then we compute the minimum number of signal power samples, denoted by n_{mn} , that we need for each representative link to ensure the required prediction accuracy for the signal powers corresponding to secondary links l_{mn} (links that exclude representative links); finally, we compute the number of samples for a representative links l_{ij} as max$\{m_{ij}, n_{mn}\}$. In what follows, we elaborate on our method of sample size computation for ensuring the required measurement accuracy and prediction accuracy.

Assume that we need to control the measurement error to be within $\beta\%$ of the mean at the $100(1-\alpha)\%$ confidence level for a representative link l_{ij}, the measured mean signal power is $\bar{p_{ij}}$, and the minimum number of samples to ensure the required accuracy is m_{ij}.

Given a set of m_{ij} samples on link l_{ij}, the $100(1-\alpha)\%$ confidence interval for the mean signal power is

$$\bar{p_{ij}} \mp z_{1-\alpha/2}\sqrt{\frac{\bar{p_{ij}}(1-\bar{p_{ij}})}{m_{ij}}} \tag{6}$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of the standard normal variate. To satisfy the required accuracy, the following should hold

$$z_{1-\alpha/2}\sqrt{\frac{\bar{p_{ij}}(1-\bar{p_{ij}})}{m}} \leq \bar{p_{ij}}\frac{\beta}{100} \tag{7}$$

From it, we have the following equation:

$$m_{ij} \geq \frac{10000 z_{1-\alpha/2}^2 \bar{p_{ij}}(1-\bar{p_{ij}})}{\beta^2} \tag{8}$$

Therefore, $\frac{10000 z_{1-\alpha/2}^2 \bar{p_{ij}}(1-\bar{p_{ij}})}{\beta^2}$ samples of packet transmission status is enough to ensure the required accuracy of $(100-\beta)\%$ for the link l_{ij}.

Given n_{ij} pairs of measurement data on link l_{ij}, we can drive the regression model $p'_{ij} = p_0 + \hat{\gamma} log(d_{ij}/d_0)$ with the corresponding standard deviation of errors $s_e = \sqrt{\frac{sum-of-squre}{n-1}}$. Here, p_0, d_0 are all predefined constant. When predicting p_{mn} for a secondary link l_{mn}, the estimating value of the predicted $\hat{p_{mn}}$ is:

$$\hat{p_{mn}} = p_0 + \hat{\gamma} log(\frac{d_{mn}}{d_0}). \tag{9}$$

Assume $x_{mn} = log(d_{mn}/d_0)$, and $x_{mn} \in \mathcal{C}_{ij}$, where $\mathcal{C}_{ij}$ is the cluster with cluster head of link l_{ij}. The standard deviation of p_{mn} is:

$$s_{\hat{p_{mn}}} = s_e[\frac{1}{n} + \frac{(x_{mn}-\bar{x})^2}{\sum_{x\in\mathcal{C}_{ij}} x^2 - n\bar{x}^2}]^{1/2} \tag{10}$$

where $\bar{x} = \frac{\sum_{x\in\mathcal{C}_{ij}} x}{|\mathcal{C}_{ij}|}$. Then the $100(1-\alpha)\%$ confidence interval for p'_{mn} is $\hat{p'_{mn}} \mp s_{\hat{p'_{mn}}} t_{[1-\alpha/2,n-2]}$, where $t_{[1-\alpha/2,n-2]}$ is the $(1-\alpha/2)$-quantile of a t-variate with $n-2$ degrees of freedom. Assume that the prediction error is required to be within $\varphi\%$ of the mean value at the $100(1-\alpha)\%$ confidence level, then the following should hold:

$$s_{\hat{p'_{mn}}} t_{[1-\alpha/2,n-2]} \leq \hat{p'_{mn}}\frac{\varphi}{100} \tag{11}$$

Then, we have the following on the required sample size n:

$$n_{ij} \geq \frac{2x_{mn}\sum_{x\in\mathcal{C}_{ij}} x - Y(\sum_{x\in\mathcal{C}_{ij}} x)^2 - \sum_{x\in\mathcal{C}_{ij}} x^2}{x_j^2 - Y\sum_{x\in\mathcal{C}_{ij}} x^2}, \tag{12}$$

where $Y = \frac{\hat{p'_{mn}}\varphi^2}{10000 s_e^2 t^2_{[1-\alpha/2,n-2]}}$. Let

$$n_{x_{mn}} = \frac{2x_{mn}\sum_{x\in\mathcal{C}_{ij}} x - Y(\sum_{x\in\mathcal{C}_{ij}} x)^2 - \sum_{x\in\mathcal{C}_{ij}} x^2}{x_{mn}^2 - Y\sum_{x\in\mathcal{C}_{ij}} x^2}. \tag{13}$$

Then the minimum required sample size:

$$n_{mn} = \max_{x\in\mathcal{C}_{ij}} n_x \tag{14}$$

VI. MEASUREMENT DISTRIBUTION

After presenting the model based and accuracy control, in this section, we answer the following question: could we use as few time as possible by carefully scheduling and distributing the measurements in the network.

The opportunity that we can take advantage of is using the simultaneous transmitting and receiving in different channels. The transmission is fully scheduled because the receiving node could only identify the source of broadcasting if they know the schedule.

We consider two kinds of measurement distribution methods, which are accuracy-constraint and time-constraint, respectively.

A. Accuracy-Constraint Method

The accuracy-constraint method requires predetermined representative links and the sample size on them. They are determined by the methods we mentioned in the previous two sections. Here, we use a vector $\{a_{ij}\}$ to represent the results, where a_{ij} is the calculated sample size in link l_{ij} and $a_{ij} = 0$ if l_{ij} is not selected as the representative link.

Thus we define the minimum cost, accuracy constraint measurement problem as:

Definition 1: Minimum Cost, Accuracy Constraint Measurement Problem: Given a network N with n nodes, each node has its own admissible band $\mathcal{M}_i$. The requirement of successful measurements is defined in $\{a_{ij}\}$. Find the schedule with earliest end time $\mathcal{T}$.

Assume that $x^m_{i,t} = 1$ denote node i will perform a broadcast in channel m, while $x^m_{i,t} = 0$. Then the maximal measurement problem could be formally defined as following optimization problem.

$$\begin{aligned} &\text{Min } \mathcal{T} \\ &\text{s.t.} \sum_{m \in \mathcal{M}_{ij}, t \leq T} (x^m_{i,t} - x^m_{i,t} x^m_{j,t}) = a_{ij}, \forall i,j \in \mathcal{N} \quad (15) \\ &\sum_{m \in \mathcal{M}_i} x^m_{i,t} = 1, \forall i \in \mathcal{N}, t \leq \mathcal{T} \quad (16) \end{aligned}$$

In our design, a successful measurement is the situation that node i broadcasts in band m while node j does not. Thus, in the formal definition, equation (15) shows that the total successful transmissions in one link is equal to the calculated measurement frequency.

Equation (16) shows that in each link we only need to select one spectrum to perform measurement, the others could be derived by the method mentioned in section IV-C. In fact, this constraint could be relaxed by changing the right part to a variable that denotes the spectrum number selected in each link to perform measurement. Usually, more spectrums mean more accuracy in deriving the signal propagation characteristic.

B. Time-Constraint Method

In some situations, the time requirement is more urgent than the accuracy requirement. Usually, we want our model as accurate as possible under a time constraint. Intuitively, we want as many collected measurements as possible. On the other hand, as the minimum estimation unit is link pairs, it is trivial that we should let the measurements be fairly distributed among all selected link pairs. Thus, we have to study the following problems.

Definition 2: Fair Measurement Problem: Given a network N with n nodes, each node has its own admissible band $\mathcal{M}_i$ and a time constraint T. The selected representative links are denoted by $\{a_{ij}\}$, where $a_{ij} = 1$ if link l_{ij} is selected and 0 otherwise. Find the schedule such that the minimum measurements of a certain link will be maximized.

This problem could be formally defined as:

$$\begin{aligned} &\text{Max } F \\ &\text{s.t. } F = \sum_{m \in \mathcal{M}_i, t \leq T} x_{i,t} + \sum_{m \in \mathcal{M}_j, t \leq T} x_{j,t} \\ &\quad - \sum_{i,j \in \mathcal{N}, m \in \mathcal{M}_{ij}, t \leq T} x^m_{i,t} x^m_{j,t} \quad (17) \\ &\sum_{m \in \mathcal{M}_{ij}, t \leq T} (x^m_{i,t} - x^m_{i,t} x^m_{j,t}) = a_{ij}, \forall i,j \in \mathcal{N} \quad (18) \\ &\sum_{m \in \mathcal{M}_i} x^m_{i,t} = 1, \forall i \in \mathcal{N}, t \leq T \quad (19) \end{aligned}$$

Both problems above are integer convex optimization problems. We can solve them approximately by relaxing the variable to a non-integer one, then rounding the result to get an integer result. The standard convex optimization procedure will not be mentioned here.

C. Summary

In summary of the aforementioned components, our solution should perform in these steps:

1) Each node collects its admissible channel information and sends to the central server.
2) On receiving all of the admissible channel information, the server first performs a schedule for all link peers to perform one successful measurement. The collected data is used to select the representative links along with the sample sizes in each link.
3) The central server will start to solve the optimization problem (in last section) according to user requirement.
4) Each node conducts broadcasting and measurement according to the predefined schedule. After grasping the measurement data, each node sends them to the central server.
5) The server uses the collected data to estimate PLE for each link and predict the receiving power. The SINR for each link are computed based on the predicted power data.

VII. Evaluation

We experimentally analyse the performance of our solution. In this section, we first present the experimental methodology and simulation settings; then we discuss the experimental results.

A. Simulation Targets

Our work mainly consists of three parts, which are, the model-based power estimation, the accuracy control and the measurement distribution. We designed the experiments to examine these three parts. Our simulation targets could be summarized as:

- The performance of the whole solution in terms of throughput improvement for the SINR-based throughput optimization algorithms.

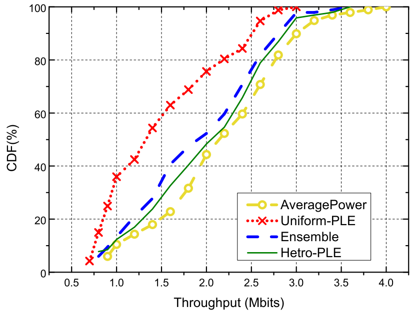

Fig. 2. The CDF of throughput under different SINR estimation methods.

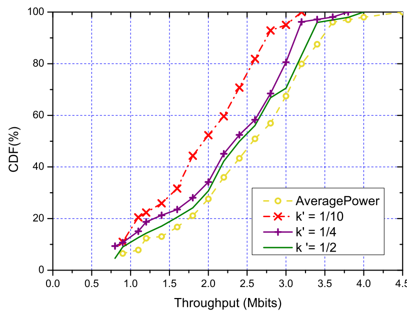

Fig. 3. The CDF of throughput with the change of the k'.

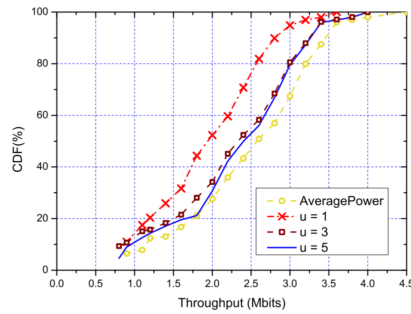

Fig. 4. The CDF of throughput with the change of the u.

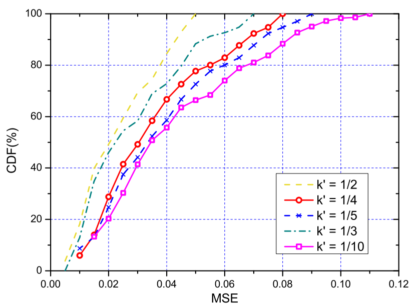

Fig. 5. The CDF of estimation MSE for different value of representative links.

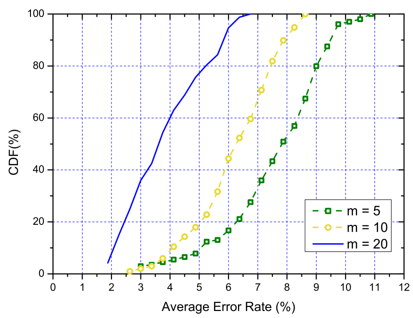

Fig. 6. The CDF of estimation error rate for different number of measurement.

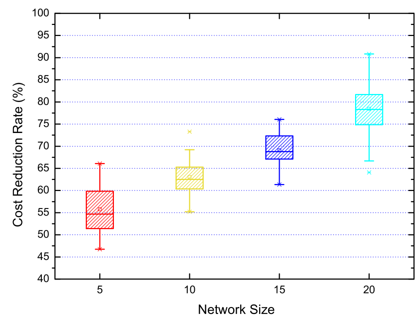

Fig. 7. The cost reduction by measurement distribution against the change of network size.

- The performance of the cross-band power propagation model should be measured in terms of how does it affect the SINR-based throughput optimization algorithms.
- The accuracy control method should be examined in terms of accuracy and overhead.
- The performance of measurement distribution should be examined in terms of average measurement time reduction.

We will design our experiments according to the above targets.

B. Simulation Settings

Our simulations are based on the data collected from Our simulations are based on the data collected from SWIM platform [18]. It is consisting of 10 nodes which is capable of running in 802.11a/b/g mode. We collect the data of the RSSI of the beacons from each AP. Generally, we activate one node at a time, while each AP will tune to 11 different channels sequentially. Then, we walk to 25 different locations (including the locations of 10 AP), to collect the 50 different beacon messages from one AP in each channel. The RSSI (Receiving Signal Strength Index), AP ID, channel ID will be recorded. In our experiment, we treat measured points as virtual APs, and use the signal strength in each beacon as the signal or interference power.

Next, we generate experimental scenarios from this data set. Generally, we extract network scenarios in this way: Our experimental scenarios are consisting of 5, 10, 15, 20 different nodes with the receiving power between them. We also set the total operating spectrum to approximately $2.4GHz$, with 11 channels of $20MHz$, which is the general settings in IEEE $802.11g$. The available channels in each node are constrained by the PU nodes. We randomly deploy a certain number of the PU nodes with assigned working channels. All the nodes within the communication range of the PU could not share the same channels. We generate 200 scenarios to perform a statistical performance comparison evaluation. The throughput of the whole network are computed using the algorithm in [6].

There are two parameters in our solution. The first is k which is the number representative links. It is related to the total link number. In order to make a fair comparison while our solution applied to different scenarios, we change this parameter into $k' = k/|\mathcal{N}|$, which is the ratio of representative links. Another parameter is the number of channel pick to perform measurement in each link, denote as u.

C. Simulation Results

1) Performance of whole solution: First, we measure the enhancement of the link scheduling and optimization algorithms brings by our solution. We conduct the algorithm in [6] with different SINR estimation methods upon all our generated scenarios. The comparison result of throughput CDF is shown in Fig. 2. In this graph, the link “AveragePower” is the result upon the SINR computed from our collected real-power data without prediction. Thus this link could serve as the optimal result. The line “Uniform-PLE” is the one with power predicted with pathloss model with uniform PLE value throughout the network. Meanwhile, line “Hetro-PLE” is the one with heterogeneous PLE value from link to link. Thus, this line could serve as the optimal result that could be achieved by pathloss model. The result of our solution is the one with mark of "Ensemble". This is performed with $k' = 1/4, u = 3$. We can see that our solution is very close to the optimal result with pathloss model. The numerical result show that it could achieve 94% of that under "Hetro-PLE" in average.

2) Performance of accuracy control: We also want to know how does the parameter k' and u affect our algorithm and the most proper value of them. Same as the examine of whole solution, we also use the throughput metric to examine the parameters. Their results are shown in Fig. 3(with $u = 2$) and Fig. 4(with $k' = 1/5$). These two parameters are linearly related to the measurement cost. We can see from these two figure that the increase of k' and u will increase the performance of our solution, while the link with $k' = 1/4$ and $u = 3$ are close enough to the optimal link.

The prediction MSE with the change of k' are shown in Fig. 5. The MSE quantities the error, when the regression model generated from representative links is used to predict the receiving power of the secondary links. We see that a small number of k' is enough to ensure small MSE.

Fig. 6 shows the CDF of average measurement errors of the representative points of all models of 16 nodes. It can be seen that the requirement of number of measurement samples is small. For instance, the average error is less than 9even with a sample size of 10. As the number of samples collected increases, the measurement error further decreases. For instance, with 20 samples, the error is usually less than 5%. From this, we see that accuracy-aware adaptive sampling can save significant sampling overhead since it only takes very few samples in general.

3) Performance of Measurement Distribution: Measurement distribution could help us schedule the measurement thus significantly reduce the cost. We measure the cost in terms of time slots. The cost is examined in scenarios with 5, 10, 15, 20 nodes, respectively. The results are shown in Fig. 7. The Y-dimension is the normalized measure of cost which is the measurement time slots over the total measurement frequency. This figure illustrates that the rate of cost reduction through measurement distribution scale up with the increase of network size. For instance, the measurement time could be saved as much as 90% when the networks consists of 20 nodes.

We can see that our solution could significantly reduce the measurement cost with tiny gap compare to the optimal result.

VIII. CONCLUSION

The SINR model, which is widely applied in scheduling and spectrum allocation, requires the power information of all the potential link peers. This information is usually assumed to be known as priori or following a uniform propagation model. Experiments show that real power data could improve the performance of the SINR-based applications with considerable margin. However, this method faces many challenges. We propose a pathloss model based solution, including a representative link selection method to cut down the measurement pairs; accuracy control to determine the sample size; and a measurement distribution method to shorten the measurement duration. Our experiments show that our solution significantly improves the SINR-based scheduling's performance.

ACKNOWLEDGEMENT

This work is supported in part by the National NSF of China under Grant No. 61073028, No. 61021062, the National Basic Research Program of China (973) under Grant No. 2009CB320705, US NSF grants ECCS 1128209, CNS 1065444, CCF 1028167, CNS 0948184 and CCF 0830289.

REFERENCES

[1] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: A survey," *Computer Networks*, vol. 47, no. 4, pp. 445–487, Mar. 2005.

[2] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer Networks*, vol. 50, no. 13, pp. 2127 – 2159, Sep. 2006.

[3] M. Andrews and M. Dinitz, "Maximizing capacity in arbitrary wireless networks in the SINR model: Complexity and game theory," in *Proceedings of IEEE INFOCOM*, 2009.

[4] O. Goussevskaia, Y. A. Oswald, and R. Wattenhofer, "Complexity in geometric SINR," in *Proceedings of ACM MobiHoc*. 2007, pp. 100–109.

[5] R. Bhatia and M. Kodialam, "On power efficient communication over multi-hop wireless networks: Joint routing, scheduling and power control," in *Proceedings of IEEE INFOCOM*, 2004.

[6] Y. Shi, T. Hou, S. Kompella, and H. D. Sherali, "Maximizing capacity in multi-hop cognitive radio networks under the sinr model," *IEEE Transactions on Mobile Computing*, vol. 99, 2010.

[7] C. Chen and D. Lee, "A joint design of distributed qos scheduling and power control for wireless networks," in *Proceedings of IEEE INFOCOM*, 2006.

[8] H. Zheng and L. Cao, "Device-centric spectrum management," in *Proceedings of the IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, 2005.

[9] L. Cao and H. Zheng, "Distributed spectrum allocation via local bargaining," in *Proceedings of IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, (SECON)*, 2005.

[10] A. T. Hoang and Y.-C. Liang, "Maximizing spectrum utilization of cognitive radio networks using channel allocation and power control," in *Proceedings of IEEE Vehicular Technology Conference. (VTC06)*, 2006.

[11] L. Cao and H. Zheng., "On the efficiency and complexity of distributed spectrum allocation," in *Proceedings of International Conference on Cognitive Radio Oriented Wireless Networks and Communications, (CrownCom)*, 2007.

[12] G. Mao, B. D. O. Anderson, and B. Fidan, "Path loss exponent estimation for wireless sensor network localization." *Computer Networks.*, vol. 51, pp. 2467–2483, July 2007.

[13] X. Zhao, L. Razoumov, and L. J. Greenstein, "Path loss estimation algorithms and results for rf sensor networks," in *Proceedings of the 61nd IEEE Vehicular Technology Conference (VTC)*, September 2004.

[14] N. Benvenuto and F. Santucci, "A least squares path-loss estimation approach to handover algorithms," *IEEE Transactions on Vehicular Technology*, vol. 48, pp. 437–447, Mar. 1999.

[15] J. Huang, S. Liu, G. Xing, H. Zhang, J. Wang, and L. Huang, "Accuracy-aware interference modeling and measurement in wireless sensor networks." in *Proceedings of the 31st Int'l Conference on Distributed Computing Systems (ICDCS)*, June 2011.

[16] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability.*, p. 281C297.

[17] "Propagation data and prediction methods for the planning of indoor radiocomm. systems and radio local area networks in the frequency range 900 mhz to 100 ghz." *Recommendation ITU-R.*, pp. 1238–1, 1999.

[18] "http://cs.nju.edu.cn/lwz/swim/swim.html."