On Quality of Monitoring for Multi-channel Wireless Infrastructure Networks

ABSTRACT

Passive monitoring utilizing distributed wireless sniffers is an effective technique to monitor activities in wireless infrastructure networks for fault diagnosis, resource management and critical path analysis. Trade-offs exist in the completeness of information captured versus the deployment and operational cost of a passive monitor system. In this paper, we introduce a quality of monitoring (QoM) metric defined by the expected number of active users monitored, and investigate the problem of maximizing QoM by judiciously assigning sniffers to channels based on knowledge of user activities in a multi-channel wireless network. Two capture models are considered. The first one, called the *user-centric model* assumes frame-level capturing capability of sniffers such that the activities of different users can be distinguished. The second one, called the *sniffer-centric model* only utilizes binary channel information (active or not) at a sniffer. For the user-centric model, we show that the implied optimization problem is NP-hard, but a constant approximation ratio can be attained via polynomial complexity algorithms. For the sniffer-centric model, we devise a stochastic inference scheme that transforms the problem into the user-centric domain, where we are able to apply our polynomial approximation algorithms. The effectiveness of our proposed scheme and algorithms is further evaluated using both synthetic data as well as real-world traces from an operational WLAN.

1. INTRODUCTION

Deployment and management of wireless infrastructure networks (WiFi, WiMax, wireless mesh networks) are often hampered by the poor visibility of PHY and MAC characteristics, and complex interactions at various layers of the protocol stacks both within a managed network and across multiple administrative domains. In addition, today's wireless usage spans a diverse set of QoS requirements from best-effort data services, to VOIP and streaming applications, making the task of managing the wireless infrastructure even more difficult, due to the additional constraints posed by QoS sensitive services. Monitoring the detailed characteristics of an operational wireless network is critical to many system administrative tasks including, e.g., fault diagnosis, resource management, and critical path analysis for infrastructure upgrades.

Passive monitoring is a technique where a dedicated set of hardware devices called *sniffers*, or monitors, are used to monitor activities in wireless networks. These devices capture transmissions of wireless devices or activities of interference sources in their vicinity, and store the information in trace files, which can be analyzed distributively or at a central location. Wireless monitoring [23, 24, 18, 6, 7] has been shown to complement wire side monitoring using SNMP and basestation logs since it reveals detailed PHY (e.g., signal strength, spectrum density) and MAC behaviors (e.g, collision, retransmissions), as well as timing information (e.g., backoff time), which are often essential for wireless diagnosis. The architecture of a canonical monitoring system consists of three components, 1) sniffer hardware, 2) sniffer coordination and data collection, and 3) data processing and mining.

Depending on the type of networks being monitored and hardware capability, sniffers may have access to different levels of information. For instance, spectrum analyzers can provide detailed time- and frequency- domain information. However, due to the limit of bandwidth or lack of hardware/software support, it may not be able to decode the captured signal to obtain frame level information on the fly. Commercial-off-the-shelf network interfaces such as WiFi cards on the other hand, can only provide frame level information[1]. The volume of raw traces in both cases tends to be quite large. Furthermore, due to the propagation characteristics of wireless signals, a single sniffer can only observe activities within its vicinity. Observations of sniffers within close proximity over the same frequency band tend to be highly correlated. Therefore, two pertinent issues need to be addressed in the design of passive monitoring systems; 1) what to monitor, and 2) how to coordinate the sniffers to maximize the amount of captured information.

This paper assumes a generic architecture of pas-

[1]Certain chip sets and device drivers allow inclusion of header fields to store a few physical layer parameters in the MAC frames. However, such implementations are generally vendor and driver dependent.

sive monitoring systems for wireless infrastructure networks, which operate over a set of contiguous or non-contiguous channels or bands[2]. To address the first question, we consider two different models for the capturing capability of the system. The first model, called the *user-centric model*, assumes availability of frame-level information such that activities of different users can be distinguished. The second model, called the *sniffer-centric model*, only assumes binary information regarding channel activities, i.e., whether *some* user is active in a specific channel near a sniffer. Clearly, the latter model imposes minimum hardware requirements, and incurs minimum cost for transferring and storing traces. We further characterize theoretically the relationship between the two models. To address the second question, we introduce a quality-of-monitoring (QoM) metric defined as the total expected number of active users detected, where a user is said to be active at time t, if it transmits over one of the wireless channels. The basic problem underlying all of our models can be cast as *finding an assignment of sniffers to channels so as to maximize the quality-of-monitoring*.

We note that the problem of sniffer assignment, in an attempt to maximize the QoM metric, is further complicated by the dynamics of real-life systems such as 1) the user population changes over time (churn), 2) activities of a single user is dynamic, and 3) connectivity between users and sniffers may vary due to changes in channel conditions or mobility. These practical considerations reveal the fundamental intertwining of "learning", where the usage pattern of wireless resources is to be estimated online based on captured information, and "decision making", where sniffer assignments are made based on available knowledge of the usage pattern, and in turn affect the QoM. In this paper, we do not address the learning problem. Rather, we focus on designing algorithms that aim at maximizing the QoM metric with different granularities of *a priori* knowledge. The usage patterns are assumed to be stationary during the decision period.

Our Contribution. In this paper, we make the following contributions toward the design of passive monitoring systems for multi-channel wireless infrastructure networks,

- We provide a formal model for evaluating the quality of monitoring.
- We study two models that differ in the capturing capability of passive monitoring systems. For each of these models we provide algorithms and methods that optimize the quality of monitoring.
- We unravel interactions between the various models.

[2]A channel can be a single frequency band, a code in CDMA systems, or a hopping sequence in frequency hopping systems.

More specifically, we show that in both the user- and sniffer-centric models considered, a pure strategy where a sniffer is assigned to a single channel suffices in order to maximize the QoM. In the *user-centric model*, we show that our problem can be formulated as a covering problem. The problem is proven to be NP-hard, and constant-approximation polynomial algorithms are provided. In the *sniffer-centric model*, we characterize the expressiveness of such a model depending on the amount of *a priori* information available. We show that somewhat surprisingly, although the only information retrieved by the sniffers in this model is binary (in terms of channel activity), it still maintains much of the "structure" of the underlying processes. Discovery of the structure using binary adaptation of the Independent Component Analysis (ICA) technique [16] allows mapping the sniffer assignment problem to the user-centric model. We complete our study by extensive evaluation using both synthetic data as well as real-world traces from an operational WLAN to demonstrate the effectiveness of our schemes in both models.

The paper is organized as follows. We first provide an overview of related work in Section 2. In Section 3, we formally introduce the QoM metric and the user-centric and sniffer-centric models for a passive monitoring system. The NP-hardness and polynomial-time algorithms for the maximum effort coverage problem which underlies two variants of the user central model are discussed in Section 4. The relationship between the user-centric and sniffer-centric models is established in Section 5, where we also describe our scheme for solving the QoM problem under the sniffer-centric. Finally, we present the results of our evaluation study using both synthetic and real traces in Section 6, and conclude the paper in Section 7.

2. RELATED WORK

Wireless monitoring is an active area of research, that has received much attention from several perspectives. There has been much work done on wireless monitoring from a *system-level* approach, in an attempt to design complete systems, and address the interactions among the components of such systems. The work in [2, 14] uses AP, SNMP logs and wired side traces to analyze WiFi traffic characteristics. Passive monitoring using multiple sniffers was first introduced by Yeo *et al.* in [23, 24], where the authors articulate the advantages and challenges posed by passive measurement techniques, and discuss a system for performing wireless monitoring with the help of multiple sniffers, synchronization and merging of the traces via broadcast beacon messages. The results obtained for these systems are mostly experimental. Rodrig *et al.* in [18] used sniffers to capture wireless data, and analyze the performance characteristics of an 802.11 WiFi network. One key contribution was the introduction of a finite state machine to infer missing frames. The Jigsaw system, that was proposed in [6], focuses on large scale monitoring using over 150 sniffers.

A number of recent works focused on the *diagnosis of wireless networks to determine causes of errors.* In [3], Chandra *et al.* proposed WiFiProfiler, a diagnostic tool that utilizes exchange of information among wireless hosts about their network settings, and the health of network connectivity. Such shared information allows inference of the root causes of connectivity problems. Building on their monitoring infrastructure, Jigsaw, Cheng *et al.* [5] developed a set of techniques for automatic characterization of outages and service degradation. They showed how sources of delay at multiple layers (physical through transport) can be reconstructed by using a combination of measurements, inference and modeling. Qiu *et al.* in [17] proposed a simulation based approach to determine sources of faults in wireless mesh networks caused by packet dropping, link congestion, external noise, and MAC misbehavior.

All the afore-mentioned work focuses on building monitoring infrastructure, and developing diagnosis techniques for wireless networks. The question of optimally allocating monitoring resources to maximize captured information remains largely untouched. In [19], Shin and Bagchi consider the selection of monitoring nodes and their associated channels for monitoring wireless mesh networks. The optimal monitoring is formulated as maximum coverage problem with group budget constraints (denoted MC-GBC), which was previously studied by Chekuri and Kumar in [4]. In this problem, we are given a ground set of n elements U, and a set S of m sets of subsets of U. Given an integer bound $d \leq m$, we are required to find a subset $T \subseteq S$ of size at most d, and for each $s \in T$ a unique subset $c_s \in s$, such that the number of elements covered by $\bigcup_{s \in T} c_s$ is maximized. One can view each element $u \in U$ as a user, each $s \in S$ as a sniffer, and each $c \in s$ as the set of users using a unique channel, that are within the sensing range of s. The problem is then to choose at most d sniffers, and assign each one to a channel, so as to maximize the number of users covered. The standard maximum coverage problem is the special case where for each $s \in S$, $|s| = 1$. Polynomial-time algorithms are devised and are shown to achieve optimal approximation ratio. The user-centric model results in a problem formulation that is similar to (albeit different from) the one addressed in [19]. On the one hand, we assume all sniffers may be used for monitoring (hence parting with our problem being akin to the classical maximum-coverage problem which is trivial for $d = m$), while on the other hand we focus on the weighted version of the problem, where elements to be covered have weights. One should note that all the lower bounds mentioned in [4, 19] do not apply to our problem.

Independent component analysis (ICA) is a computational method for separating a multivariate signal into additive subcomponents supposing the mutual statistical independence of the non-Gaussian source signals. Most ICA methods assume linear mixing of continuous signals [16]. A special variant of ICA, called binary ICA (BICA), considers boolean mixing (e.g., OR, XOR etc.) of binary signals, and has been applied in the context of multi-assignment clustering for boolean data [22], and medical diagnosis [8], etc. Existing solutions to BICA mainly differ in their assumptions of prior distribution of the mixing matrix, noise model, and/or hidden causes. For instance, in [8], infinite number of hidden causes following the same Bernoulli distribution are assumed. Accordingly, reversible jump Markov chain Monte Carlo and Gibbs sampler techniques are applied. In contrast, in our sniffer-centric model, the hidden causes may follow different distribution and the mixing matrix tends to be sparse.

3. PROBLEM FORMULATION

3.1 Network model and QoM metric

Consider a system of m sniffers, and n users, where each user u operates in one of K channels, $c(u) \in \mathcal{K} = \{1, \ldots, K\}$. The users can be wireless (mesh) routers, access points or mobile users. At any point in time, a sniffer can only monitor packet transmissions over a single channel. We assume the propagation characteristics of all channels are similar. We represent the relationship between users and sniffers using an undirected bipartite graph $G = (S, U, E)$, where S is the set of sniffer nodes and U is the set of users. An edge $e = (s, u)$ exists between sniffer $s \in S$ and user $u \in U$ if s can capture the transmission from u. If transmissions from a user cannot be captured by any sniffer, the user is excluded from G. For every vertex $v \in S \cup U$, we let $N(v)$ denote vertex v's neighbors in G. For users, their neighbors are sniffers, and vice versa. Abusing the notation slightly, we also refer to G as the binary adjacency matrix of graph G.

We will consider *sniffer assignments* of sniffers to channels, $a : S \to \mathcal{K}$. Given a sniffer assignment a, we consider a partitioning of the set of sniffers $S = \bigcup_{k=1}^{K} S_k$, where S_k is the set of sniffers assigned to channel k. We further consider the corresponding partition of the set of users $U = \bigcup_{k=1}^{K} U_k$, where U_k is the set of users operating in channel k. Let $G_k = (S_k, U_k, E_k)$ denote the bipartite subgraph of G induced by channel k. Given any sniffer s, we let $N_k(s) = N(s) \cap U_k$, i.e., the set of neighboring users of s that use channel k.

A *monitoring strategy* determines the channel(s) a sniffer monitors. It could be a *pure strategy*, i.e., the channel a sniffer is assigned to is fixed, or a mixed strategy where sniffers choose their assigned channel in each slot according to a certain distribution. Formally, let $\mathcal{A} = \{a \mid a : S \to \mathcal{K}\}$ be the set of all possible assignments. Let $\pi : \mathcal{A} \to [0, 1]$ be a probability distribution over the set of sniffer assignments. We refer to such a distribution as a *mixed strategy.* Clearly, a pure strategy a that selects a single channel per sniffer is a special case of mixed strategies, namely, $\pi(a) = 1$.

We measure the quality of monitoring (QoM) by the total expected number of active users monitored by the set of sniffers. This measure is called the *QoM metric.*

3.2 Models for Observing User Access Patterns

In this section, two parametric models are proposed to describe the observability of usage patterns. We assume time is slotted, and that all channel and users' statistics remain stationary for a period of time T.

User-centric model. First, we consider transmission events in the network from the user's viewpoint. We assume that the bipartite graph G is known by inspecting the packet header information from each sniffer's captured traces.

In the user-centric model, the transmission probabilities of the users $\overline{p} = (p_u)_{u \in U}$ are known and assumed to be independent [3] (p_u denotes the transmission probability of user u)[4]. Clearly, if p_u is set to 1, user u always has packets to transmit. This can be used to model scenarios where worst-case traffic load is assumed.

Sniffer-centric model. The user-centric model requires detailed knowledge of each user's activities. This necessitates frame-level capturing capability by the passive monitoring system. In the sniffer-centric model, only **binary** information (*on* or *off*) of the channel activity at each sniffer is assumed.

Let $\mathbf{x} = [x_1, x_2, \ldots, x_m]$ be a vector of m binary random variables, where x_i denotes whether or not sniffer s_i captures communication activities in its associated channel. We further denote by $\mathbf{x}_k$ the binary vector of observations for sniffers S_k (operating on channel k). We will sometimes abuse notation and let $\mathbf{x} = \{\mathbf{x}_k \mid k = 1, \ldots, K\}$. We assume that sniffers' observations in different channels are independent. However, dependency exists among observations of sniffers operating in the same channel (as a result of transmissions made by the same set of users). Given an assignment a, a complete characterization of the sniffers' observations is given by the joint probability distribution $\mathcal{P}_a(\mathbf{x}_k)$, $k = 1, \ldots, K$. By independence of different channels we have $\mathcal{P}_a(\mathbf{x}) = \prod_{k=1}^{K} \mathcal{P}_a(\mathbf{x}_k)$.

Clearly, the sniffer-centric model is not as expressive as the user-centric model (formally proven in Section 5.1). However, it has the advantage of being based on *aggregated* statistics, which are likely to remain stationary in the presence of moderate user-level dynamics, such as joining and leaving the networks, or changes in transmission activities (e.g., busy or thinking time). Furthermore, obtaining such binary information is less costly in both hardware requirements and communication/storage complexity.

[3]Making the assumption that user activities are independent are widely adopted in literature, examples are [10] and [20].

[4]The proposed optimization framework is valid for both IID and non-IID user processes.

4. QOM UNDER THE USER-CENTRIC MODEL

Under the user-centric model, the goal is to maximize the expected number of active users monitored. Recall that p_u is the transmission probability of user u. This problem can be formulated formally by:

$$\begin{array}{ll} \max & \sum_{u \in U} p_u \times \sum_{a \in \mathcal{A}(u)} \pi(a) \\ \text{s.t.} & \pi(a) \in [0, 1] \\ & \sum_{a \in \mathcal{A}} \pi(a) = 1, \end{array} \tag{1}$$

where

$$\mathcal{A}(u) = \{a \mid \exists s \in N(u) \text{ s.t. } a(s) = c(u)\},$$

i.e., $\mathcal{A}(u)$ is the set of assignments that monitors user u.

The objective function can be rewritten as

$$\sum_{a \in \mathcal{A}} \pi_a \sum_{u \in U} p_u \cdot \mathbb{1}(a \in \mathcal{A}(u)), \tag{2}$$

where $\mathbb{1}(\cdot)$ is an indicator function. From Eq. (2) it is clear that a pure strategy can be adopted, i.e., an optimal assignment is given by

$$a^* = \arg\max \sum_{u \in U} p_u \cdot \mathbb{1}(a \in \mathcal{A}(u)).$$

4.1 Problem formulation

Let MAX-EFFORT-COVER (MEC) denote the problem of finding the largest (weight) set of users that can be monitored by a set of sniffers, where each sniffer can monitor one of a set of k channels. Note that in MEC the weights can in fact be any non-negative values and are not limited to $[0, 1]$. The MEC problem can be cast as the following integer program (IP):

$$\begin{array}{lll} \max & \sum_{u \in U} p_u y_u & \\ \text{s.t.} & \sum_{k=1}^{K} z_{s,k} \leq 1 & \forall s \in S \\ & y_u \leq \sum_{s \in N(u)} z_{s,c(u)} & \forall u \in U \\ & y_u \leq 1 & \forall u \in U \\ & y_u, z_{s,k} \in \{0, 1\} & \forall u, s, k. \end{array} \tag{3}$$

Each sniffer is associated with a set of binary decision variables, $z_{s,k} = 1$ if the sniffer is assigned to channel k; 0, otherwise. y_u is a binary variable indicating whether or not user u is monitored, and p_u is the weight associated with user u.

One should first note that the problem is trivial if $k = 1$, since all sniffers would simply be assigned to the sole available channel. We can therefore assume that $k \geq 2$.

The MEC problem can be viewed as a special case of the MC-GBC (described in Section 2), where we have $d = m$. One should note that previous hardness results for MC-GBC (both NP-hardness, as well as hardness of approximation) were based on a reduction to the standard maximum coverage problem for $d < m$ (the maximum coverage problem becomes trivial for $d = m$). It follows that none of these proofs are applicable to the MEC problem. Surprisingly, there has not been any work done explicitly on the MEC problem, which seems to be a natural and important variant of the maximum coverage problem.

4.2 Hardness of MEC

In what follows we show that the MEC problem is NP-hard for $k \geq 2$, even for the unweighted case (i.e., where $p_u = 1$ for all $u \in U$). Our proof shows that the hardness of the MEC problem actually follows from the choices available to the different sniffers. It is inherently different from the hardness suggested for the MC-GBC problem, which follows from limiting the number of sniffers one is allowed to use.

The proof uses a reduction from MONOTONE-3SAT, which is known to be NP-hard (see [12, 11]). In MONOTONE-3SAT we are given as input an instance of 3SAT where every clause consists of either solely positive variables, or solely negated variables. The goal is to decide whether or not there exists an assignment which satisfies all clauses.

THEOREM 1. *The unweighted* MEC *problem is NP-hard, even for $k = 2$.*

PROOF. Let $C = \{C_1, \ldots, C_n\}$ be a set of 3SAT clauses, where each clause C_i is either the disjunction of 3 positive variables, or the disjunction of 3 negated variables, over a ground set of variables $X = \{x_1, \ldots, x_m\}$. We construct the following instance to MEC with $k = 2$ channels: for every variable x_i we define a sniffer s_i, and for every clause C_j we define a user u_j. We let $c(u_j) = 1$ if C_j consists solely of positive variables, and $c(u_j) = 2$ if C_j consists solely of negated variables (note that this definition is consistent since we start with an instance of MONOTONE-3SAT, where in every clause all variables agree in sign). We now define the bipartite graph $G = (U, S, E)$ which defines which user is in the range of which sniffer. We do this by defining the neighboring users of every sniffer and every channel. For every $i = 1, \ldots, m$ we define

$$N(s_i) = \{u_j \mid x_i \in C_j\} \cup \{u_j \mid \neg x_i \in C_j\}$$

The set of edges E is therefore defined by

$$E = \{(u_j, s_i) \mid u_j \in N(s_i)\}.$$

Given a channel assignment $a : S \to \{1, \ldots, K\}$, we define a truth assignment ϕ for the variables in X as follows:

$$\phi(x_i) = \begin{cases} T & a(s_i) = 1 \\ F & a(s_i) = 2 \end{cases}$$

Clearly ϕ is well defined. We now show that assignment a is able to monitor all the users if and only if all the clauses are satisfied by truth assignment ϕ. Assume a is able to monitor all the users, and let C_j be a clause in C. By the assumption, u_j must be monitored by at least one sniffer s_i such that $a(s_i) = c(u_j)$ and $u_j \in N(s_i)$. Assume C_j consists of solely positive variables. It follows that s_i is assigned to channel $c(u_j) = 1$. It follows that $\phi(x_i) = T$. Since by the definition of $N(s_i)$ we have that $x_i \in C_j$ (since C_j is a clause of solely positive variables), we are guaranteed that truth assignment ϕ satisfies C_j. The case where C_j consists of solely negative variables is symmetric. Assume now that a does not monitor all the users, and let u_j be an un-monitored user. Assume C_j consists of solely negative variables, which in turn implies that $c(u_j) = 2$. Since u_j is unmonitored, it follows that for every sniffer s_i such that $u_j \in N(s_i)$, $a(s_i) = 1$. By the definition of the reduction, and the definition of ϕ, this implies that for every variable x_i which appears in C_j, $\phi(x_i) = T$. Since all these variables appear in their negated form in C_j, C_j is not satisfied by assignment ϕ. Again, the case where C_j consists of solely positive variables is symmetric. □

Theorem 1 implies that one would have to settle for approximate solutions to MEC. We first note that Guruswami and Khot show in [13] that MONOTONE-3SAT is NP-hard to approximate within a factor of $7/8+\varepsilon$ for every $\varepsilon > 0$. The following is a corollary of the above fact, and the proof of Theorem 1:

COROLLARY 2. *The* MEC *problem is NP-hard to approximate to within a factor of $7/8 + \varepsilon$ for every $\varepsilon > 0$.*

PROOF. By closely examining the reduction appearing in the proof of Theorem 1, it follows that the number of satisfied clauses given the truth assignment ϕ is exactly the number of users that are covered by the channel assignment a. It therefore follows that the reduction is approximation-preserving (i.e., any α-approximation for MEC implies an α-approximation for MONOTONE-3SAT). Combining this with the fact that it is NP-hard to approximate MONOTONE-3SAT to within a factor of $7/8 + \varepsilon$ for every $\varepsilon > 0$ ([13]), the result follows. □

4.3 Algorithms for MEC

As previously stated, since MEC is a special case of the MC-GBC problem, we can use the available approximation algorithms for MC-GBC (e.g., [4, 19]) to solve our problem in the user-centric model. In what follows we give a brief overview of the algorithms we use.

These algorithms would serve as a crucial component in maximizing QoM in the more oblivious settings of the sniffer-centric model (where no *a priori* knowledge of the problem's structure is available), as we discuss in Section 5.

The greedy algorithm. The greedy algorithm GREEDY iteratively assigns sniffers to users, where at each step it chooses the sniffer and the assignment that (locally) maximizes the weight of coverage of those not yet monitored users.

It is proven in [4] that in the unweighted case, i.e., where all users have the same weight, GREEDY guarantees to produce a $\frac{1}{2}$-approximate solution, and that this is tight. The following theorem shows that the same holds also for the weighted case, which generalizes the MEC problem.

THEOREM 3. GREEDY *is a* $\frac{1}{2}$*-approximation algorithm for the weighted* MC-GBC *problem.*

PROOF. The proof follows closely the proof appearing in [4], and is omitted due to space limit. □

Note that the example provided in [4] showing that this analysis is tight naturally also holds for the weighted case.

LP-based algorithm. This algorithm is based on solving the LP-relaxation of the IP formulation for MEC appearing in Eq. (3). Once we have an optimal solution to the LP-relaxation, we round the fractional solution into an integral solution, with e.g., the probabilistic rounding technique of Srinivasan [21]. We next sketch the basic idea of this probabilistic rounding technique. Let z^* be an optimal solution to the LP relaxation of Eq. (3), and let s be any sniffer. If $\sum_k z^*_{s,c} > 0$, one can view the induced solution $z^*_s : C \to [0,1]$ as a probability measure over the different channels (via normalization). The goal is to decide on an integral channel assignment for s, namely, setting each $z^*_{s,c}$ to a value in $\{0,1\}$ such that *exactly* one variable out of the k variables corresponding to sniffer s is set to the value 1. The algorithm builds a binary tree whose leaves corresponds to the k variables $z_{s,k}$ associated with sniffer s, and pairs unset variables in a bottom-up fashion. The pairing is made such that an internal node sets at least one of the variables corresponding to its children. This is done while adjusting the (probability) value of the (other) unset variable. This approach is proven to produce a valid assignment in linear time [21]. We refer to the above algorithm as PROBRAND.

As mentioned in [4] (and later made explicit in [19]), this method produces a $(1-1/e)$-approximate solution, for the case where all users have the same weight. As is the case with the greedy algorithm, the analysis for the unweighted case (e.g., appearing in [19]) can be extended to provide the same guarantee also for the weighted case, as demonstrated by the following theorem (the proof is omitted due to space constraints).

THEOREM 4. PROBRAND *is a* $(1-1/e)$*-approximation algorithm for the weighted* MC-GBC *problem.*

One could also use the pipage LP-based technique suggested by Ageev and Sviridenko [1], as an alternative to PROBRAND. This approach has the exact same approximation guarantee as PROBRAND.

We note that the approximation guarantee of the LP-based algorithms are best possible for the MC-GBC problem, again, due to a reduction from maximum coverage, and the fact that it is hard to approximate the maximum coverage problem to within a factor better than $(1-1/e)$ [9]. However, this lower bound does not necessarily hold for the MEC problem, for which the previous lower bound of $7/8+\varepsilon$ for every $\varepsilon > 0$ is the best available.

5. QOM UNDER THE SNIFFER-CENTRIC MODEL

Recall that in the sniffer-centric model, given an assignment $a \in \mathcal{A}$, $\prod_{k=1}^{K} \mathcal{P}_a(\mathbf{x}_k)$ is the probability distribution of binary observations $\mathbf{x} = \{\mathbf{x}_k, k = 1, \ldots, K\}$ from m sniffers. Let $w(\mathbf{x}_k)$ be the expected number of active users monitored by sniffers in channel k. The MEC problem under the sniffer-centric model is defined as follows.

$$\begin{array}{ll} \max & \sum_{a\in\mathcal{A}} \pi(a) \sum_{k=1}^{K} w(\mathbf{x}_k)\mathcal{P}_a(\mathbf{x}_k) \\ \text{s.t.} & \pi(a) \in [0,1] \\ & \sum_{a\in\mathcal{A}} \pi(a) = 1, \end{array}$$

Clearly, a pure strategy suffices, i.e., there exists an optimal assignment such that,

$$a^* = \arg\max \sum_{k=1}^{K} w(\mathbf{x}_k)\mathcal{P}_a(\mathbf{x}_k). \quad (4)$$

Note that in contrast to the user-centric model, where transmission activities from different users are independent, observations of sniffers are correlated. As a result, $\mathcal{P}_a(\mathbf{x}_k)$ cannot be simplified as a product form. This motivates us to exploit the underlying (though not directly observable) independence among users, and map the problem to QoM under the user-centric model.

5.1 Relationship between the user-centric and sniffer-centric models with known G and unknown $\bar{p}$

In the sniffer-centric model, each sniffer only reports binary output regarding the channel activities, and thus the access probability of the users as well as the bipartite graph G, are both *hidden*. In this section we derive the sufficient and necessary conditions for unraveling the access probabilities of the users given G and $\mathcal{P}(\mathbf{x})$.

Let $\mathbf{y} = [y_1, y_2, \ldots, y_n]$ be a vector of n binary random variables, where $y_j = 1$ if user u_j transmits in its associated channel, and $y_j = 0$ otherwise. $\mathbf{y}_c$ is the vector of activities for users transmitting in channel k (i.e., users in U_k). The joint distribution of $\mathbf{y}$ is given by $\mathcal{P}(\mathbf{y}) = \prod_{y_j=1} p_j \prod_{y_j=0}(1 - p_j)$. The product form is due to the independence among users' activities. The main question we aim to answer is the following: Given the vector $\mathbf{x}_k$ of sniffers' observations, what knowledge can be obtained regarding $\mathbf{y}_c$? Throughout this section, unless otherwise specified, we limit the discussion to users and sniffers in a fixed channel k, and drop the subscript. We will also denote by g_{ij} the entry in the i'th row and j'th column of G. Using the adjacency matrix, we have the following,

$$x_i = \bigvee_{j=1}^{n} g_{ij} \wedge y_j, \; i = 1, \ldots, I, \tag{5}$$

where $\wedge$ is Boolean *AND* and $\vee$ Boolean *OR*. Define the set

$$Y(\mathbf{x}) = \{\mathbf{y} \mid \bigvee_{j=1}^{n} g_{ij} \wedge y_j = x_i, \forall i\}.$$

Therefore,

$$\mathcal{P}(\mathbf{x}) = \mathcal{P}(\mathbf{y} \in Y(\mathbf{x})) = \sum_{\mathbf{y} \in Y(\mathbf{x})} \mathcal{P}(\mathbf{y}) \tag{6}$$

The necessary and sufficient conditions that uniquely determining $\overline{p}$ using G and $\mathcal{P}(\mathbf{x})$ is characterized in the following theorem.

THEOREM 5. *Given $G = (S, U, E)$, $\overline{p}$ can be uniquely determined by $\mathcal{P}(\mathbf{x})$ iff $\forall u_j \neq u_{j'} \in U$, $N(u_j) \neq N(u_{j'})$.*

PROOF. The proof is omitted due to space limit. □

The above theorem essentially shows that in the sniffer-centric model, if the adjacency matrix is known, then one can effectively determine the transmission probabilities of the users from the joint distribution of sniffers' observations. In presence of measurement noise, methods such as Expectation-Maximization can be applied. We therefore obtain an instance of the problem corresponding to our user-centric model, which can be solved efficiently using the algorithms described in Section 4.3.

5.2 Inference of G and unknown $\overline{p}$ using binary ICA

In this section we deal with the sniffer-centric model, and the problem of QoM in a scenario where both the user access probabilities and the adjacency matrix are unknown. In such a scenario, the question is whether knowledge regarding G and $\overline{p}$ can still be obtained from $\mathcal{P}(\mathbf{x})$. The answer is positive. Consider the simple example, where two sniffers s_1 and s_2 observe the activity of a single user u. In this case, $x_1 = x_2$. Therefore,

$$\begin{aligned} \mathcal{P}(\mathbf{x}) &= \mathcal{P}(x_1)\mathbb{1}(x_1 = x_2) \\ &= \mathcal{P}(y_u = x_1) \\ &= \begin{cases} p_u, & x_1 = 1 \\ 1 - p_u, & x_1 = 0 \end{cases}. \end{aligned}$$

Therefore, if the joint distribution of $\mathbf{x}_k$ is the product of a marginal distribution with an indicator function, and the two marginal distributions are identical, we can infer that both sniffers observe the same set of users. Generally, the joint distribution of $\mathbf{x}$ preserves a certain stochastic "structure" of the user's activities. We will formalize this observation in the subsequent section by devising an inference method to estimate G and $\overline{p}$ from $\mathcal{P}(\mathbf{x})$.

Estimation of G. This problem is similar to the problem addressed by the Independent Component Analysis (ICA) scheme [16], where the observed data is expressed as a linear transformation of latent variables that are non-Gaussian and mutually independent. ICA is widely used in applications such as blind source separation, feature extraction, noise reduction in imaging data etc. In the classic ICA model, continuous-value variables are mixed linearly:

$$\mathbf{x} = G\mathbf{y},$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_m)^T$ is the vector of observed random variables, $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T$ is the vector of independent latent variables (the "independent components"), and G is an unknown constant matrix, called the mixing matrix. The problem is then to estimate both the mixing matrix G and the distribution of the latent variables y_i, using observations of $\mathbf{x}$ alone. ICA can be solved by casting it as an optimization problem which aims at maximizing the nongaussianity of estimates $\hat{\mathbf{y}}$ (thus aiming at preserving the level of independence in $\mathbf{y}$). (see [16]).

ICA assumes that both $\mathbf{y}$ and $\mathbf{x}$ are continuous random variables and linear mixing of $\mathbf{y}$, and thus is not directly applicable to our problem. In Eq. (5), $\mathbf{x}$ and $\mathbf{y}$ are binary random variables, and Boolean operations are used. In [15] Himberg *et al.* show that with a proper transformation, extensions of well-known ICA algorithms work well when the data is sparse enough. We adopt the algorithm presented in [15] with some modifications. The basic idea is as follows.

First, Eq. (5) is simplified using linear mixing and a (coordinate-wise) unit step function.

$$\mathbf{x} = U(G\mathbf{y}), \tag{7}$$

where $U(\cdot)$ is unit step functions defined by $U(r) = \mathbb{1}(r > 0)$.

By applying the standard ICA, an estimation of the *linear* mixing matrix $\hat{G}_L$ can be found by ignoring the step function. Then, the *binary* mixing matrix $\hat{G}$ can

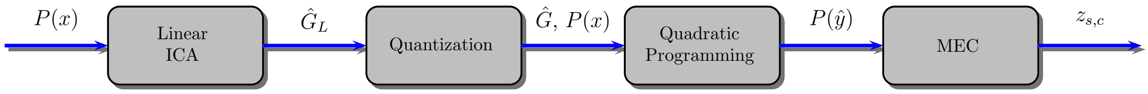


Figure 1: Channel selection algorithm under the sniffer-centric model

be estimated by a *quantization* operation, defined by

$$\hat{G} = U(\Lambda^{-1}\hat{G}_L - \boldsymbol{T}). \tag{8}$$

The diagonal scaling matrix Λ has

$$\lambda_{ii} = \text{signmax}(\hat{g_L}_i),$$

where

$$\text{signmax}(\boldsymbol{r}) = \begin{cases} \max(\boldsymbol{r}) & \text{if } |\max(\boldsymbol{r})| > |\min(\boldsymbol{r})| \\ \min(\boldsymbol{r}) & \text{otherwise.} \end{cases}$$

Λ scales the elements in the mixing matrix to the maximum value 1. The matrix $\boldsymbol{T}$ contains thresholds, such that the higher the threshold value, the sparser $\hat{G}$ is. We note that our definition of the signmax is different than the one used in [15].[5]

Estimation of $\mathcal{P}(\mathbf{y})$. Once $\hat{G}$ is determined, $\mathcal{P}(\mathbf{y})$ needs to be estimated. From $x_i = U(\hat{g}_i y_i)$, where $\hat{g}_i$ is the ith row of $\hat{G}$, we have,

$$p(x_i = 0) = \prod_{\hat{g}_{ij}=1} p(y_j = 0).$$

The product is due to independence of y_i's. Taking $\log(\cdot)$ on both sides, we have

$$\log(p(x_i = 0)) = \sum_{\hat{g}_{ij}=1} \log(p(y_j = 0)).$$

Let $\alpha_i = \log(p(x_i = 0))$, and $\beta_i = \log(p(y_j = 0))$. Define $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_m]^T$, and $\beta = [\beta_1, \beta_2, \ldots, \beta_n]^T$. Therefore, we formulate the following optimization problem.

$$\begin{aligned} \min \quad & \| \alpha - \hat{G}\beta \|^2 \\ \text{s.t} \quad & \beta < 0, \end{aligned} \tag{9}$$

where $\| \cdot \|$ is the norm of a vector. Clearly, this is a constrained quadratic programming problem with a positive semi-definite matrix (i.e., all eigenvalues are non-negative), and can be solved in polynomial time.

Channel selection. With the estimates $\hat{\mathbf{y}}$ and $\hat{G}$ at hand, we effectively transform the sniffer-centric model to the user-centric model. The method described in Section 4.3 can then be applied to determine the channel assignment of each sniffer. The complete channel assignment scheme in the sniffer-centric model is illustrated in Figure 1.

[5]We believe the form of signmax in [15] is not correct.

6. EVALUATION

In this section we evaluate the performance of different algorithms under the user-centric and sniffer-centric models using both synthetic and real traces. Synthetic traces allow us to control the parameter settings by varying the number of users, the number of channels as well as the traffic load of users, and investigate their effects on the performance of different algorithms. The real-world traces are collected from an operational WLAN. They provide insights on the performance under realistic traffic loads and user distributions over space.

In addition to the greedy and LP-based algorithm, two baseline algorithms are considered:

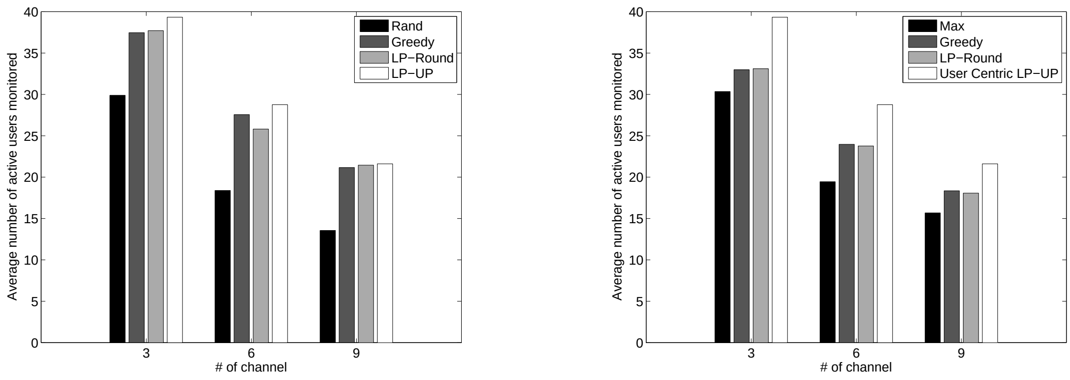
- **Random channel assignment (Rand)**, where the sniffer channels are assigned randomly. Rand assumes no prior information regarding either user activties or snifffers' observations. It is independent of whether the user or sniffer centric model is assumed. However, for conciseness of presentation, we only compare it side-by-side with results of user-centric models.
- **Max Sniffer Channel (Max)**, where a sniffer is assigned to its busiest channel. This scheme is the most intuitive approach in the sniffer-centric model when sniffers decide their channel assignment *non-cooperatively* based on local observations. Note it is easy to construct scenarios where Max performs arbitrariliy bad. Thus, its worst case performance is unbounded.

For the inference scheme in the sniffer-centric model, we used the FastICA algorithm [16] to compute the linear mixing matrix $\hat{G}_L$.

6.1 Synthetic traces

In this set of simulations, 1000 wireless users are placed randomly in a 500x500 square meter area. The area is partitioned into hexagon cells with circumcircle of radius 86 meters. Each cell is associated with a base station operating in a channel (and so are the users in the cell). The channel to base station assignment ensures that *no neighboring cells use the same channel*. 25 Sniffers are deployed in a grid formation separated by distance 100 meters, with a coverage radius of 120 meters. uniformly from [0, 0.06], resulting in an average busy probability of 0.2685 in each cell. We vary the total number of orthogonal channels from 3 to 9[6]. The

[6]In 802.11a networks, there are 8 orthogonal channels in 5.18-5.4GHz, and one in 5.75GHz.



(a) User-centric model

(b) Sniffer-centric model

Figure 2: QoM under user-centric and sniffer-centric models with synthetic traces

results shown are the average of 20 runs with different seeds.

Figure 2(a) shows the simulation results under the user-centric model. One can see that the performance of the greedy algorithm and LP-based algorithm with random rounding are comparable to the LP upper bound. When the number of channels is small, a random assignment can yield a reasonable monitoring quality. However, as the number of channels increases, some channels of a sniffer do not have any activity leading to poorer performance of the random assignment.

Figure 2(b) summarizes the simulation results for the sniffer-centric model. In this part, we compare the QoM using the Max algorithm, with that attained by the greedy and LP-based channel assignment, applied to the mixing matrix $\hat{G}$ and user access probabilities $\overline{p}$ inferred by our scheme depicted in Figure 1. We observe that the algorithms based on the inferred $\hat{G}$ and $\hat{y}$ outperform the Max algorithm. Recall that according to the Max algorithm, a sniffer non-cooperatively decides its own channel assignment and selects the most active channel. Clearly, the Max algorithm does not take into account the correlations among the observations of neighboring sniffers in the same channel. In contrast, the proposed inference algorithm can indeed extract such a correlative structure from the binary observations. We further note that by comparing Figure 2(a) and Figure 2(b), one can detect QoM degradations due to uncertainty in G and $\overline{p}$. Such degradations are expected due to the information loss in having merely the binary observations of the sniffers.

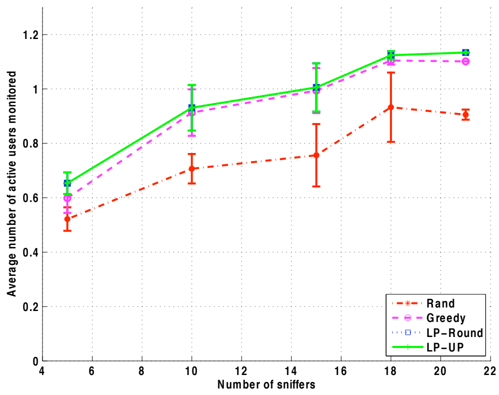
6.2 Real traces

In this section, we evaluate our proposed schemes using real traces collected from the campus wireless network using 21 WiFi sniffers deployed in our building. Over a period of 6 hours, between 12pm and 6pm, each sniffer captured approximately 300,000 MAC frames. Altogether, 655 unique users are observed operating over three channels. The number of users observed on channels 1, 2, 3 are 382, 118, and 155, respectively. The average active probability of a user is around 0.0014.

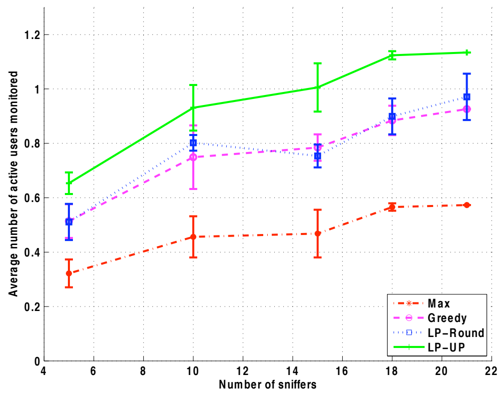
Figure 3 gives the average number of active users monitored under the user-centric and sniffer-centric models. The number of sniffers in the experiments varies from 5 to 21 by including only traces from the corresponding sniffers. The number of channels is fixed at 3. Except for the case with 21 sniffers, all data points are averages of 5 scenarios with different sets of sniffers, chosen uniformly at random. Recall that the average active probability is 0.0014. Thus, the number of users active in a slot is around 1. In the user-centric cases (Figure 3(a)), both the greedy and the LP-round algorithms with random rounding significantly outperform the random assignment. Moreover, their performance is comparably with the LP upper bound on the optimal performance possible. As the number of sniffers increases, the average number of users monitored increases but tends to flatten out since most users have been monitored. The gap between greedy/LP-round and random also reduces from 31.6% at 5 sniffers to about 10% at 21 sniffers. In the sniffer-centric case, both the greedy and the LP-round algorithms outperform Max. However, we do observe performance degradation due to the uncertainty in G and $\overline{p}$, as would be expected due to the loss of information, when compared to the performance in the user-centric model.

7. CONCLUSION

In this paper, we formulate the problem of maximizing QoM in multi-channel infrastructure wireless networks with different *a priori* knowledge. Two different models are considered, which differ by the amount (and type) of information available to the sniffers. We show that when complete information of the underlying cover graph and the access probabilities of users is available, the problem is NP-hard, but can be approximated within a constant factor. We further show that



(a) User-centric model



(b) Sniffer-centric model

Figure 3: QoM under user-centric and sniffer-centric models with real WiFi traces. In the user-centric model, the results of LP-round coincide with that of the LP-UP. In some cases, the confidence interval is quite small and is thus not observable in the figures.

when only binary information about channel activities is available to the sniffers, one can map the problem to the the one where complete information is at hand using the statistics of the sniffers' observations. Evaluations demonstrate the effectiveness of our proposed inference method and optimization techniques.

There are a few fundamental open questions that remain to be addressed, including 1) design and analysis of binary inference schemes, with provable performance bounds, and 2) resolving the gap between the upper and lower bounds on the approximation ratio of the MEC problem.

8. REFERENCES

[1] A. A. Ageev and M. Sviridenko. Pipage rounding: A new method of constructing algorithms with proven performance guarantee. *Journal of Combinatorial Optimization*, 8(3):307–328, 2004.

[2] A. Balachandran, G. M. Voelker, P. Bahl, and P. V. Rangan. Characterizing user behavior and network performance in a public wireless LAN. *SIGMETRICS Perform. Eval. Rev.*, 30(1):195–205, 2002.

[3] R. Chandra, V. N. Padmanabhan, and M. Zhang. WiFiProfiler: cooperative diagnosis in wireless LANs. In *MobiSys*, pages 205–219, 2006.

[4] C. Chekuri and A. Kumar. Maximum coverage problem with group budget constraints and applications. In *APPROX*, pages 72–83, 2004.

[5] Y.-C. Cheng, M. Afanasyev, P. Verkaik, P. Benkö, J. Chiang, A. C. Snoeren, S. Savage, and G. M. Voelker. Automating cross-layer diagnosis of enterprise wireless networks. In *SIGCOMM*, 2007.

[6] Y.-C. Cheng, J. Bellardo, P. Benkö, A. C. Snoeren, G. M. Voelker, and S. Savage. Jigsaw: solving the puzzle of enterprise 802.11 analysis. In *SIGCOMM*, 2006.

[7] A. Chhetri and R. Zheng. WiserAnalyzer: A passive monitoring framework for wlans. In *Proceedings of the 5th International Conference on Mobile Ad-hoc and Sensor Networks (MSN)*, 2009.

[8] F. W. Computer and F. Wood. A non-parametric bayesian method for inferring hidden causes. In *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence (UAI*, pages 536–543. AUAI Press, 2006.

[9] U. Feige. A threshold of $\ln n$ for approximating set cover. *Journal of the ACM*, 45(4):634–652, 1998.

[10] A. A. Garba, R. M. H. Yim, J. Bajcsy, and L. R. Chen. Analysis of optical cdma signal transmission: capacity limits and simulation results. *EURASIP J. Appl. Signal Process.*, 2005:1603–1616, 2005.

[11] M. R. Garey and D. S. Johnson. *Computers and Intractability - A Guide to the Theory of NP-Completeness.* W.H. Freeman and Company, New York, 1979.

[12] E. M. Gold. Complexity of automaton identification from given data. *Information and Control*, 37(3):302–320, 1978.

[13] V. Guruswami and S. Khot. Hardness of Max 3SAT with no mixed clauses. In *Proceedings of the 20th Annual IEEE Conference on Computational Complexity (CCC)*, pages 154–162, 2005.

[14] T. Henderson, D. Kotz, and I. Abyzov. The changing usage of a mature campus-wide wireless network. In *Mobicom*, pages 187–201, 2004.

[15] J. Himberg and A. Hyvrinen. Independent component analysis for binary data: An experimental study. In *Int. Workshop on Independent Component Analysis and Blind Signal Separation (ICA2001)*, 2001.

[16] A. Hyvärinen and E. Oja. Independent component analysis: algorithms and applications. *Neural Netw.*, 13(4-5):411–430, 2000.

[17] L. Qiu, P. Bahl, A. Rao, and L. Zhou. Troubleshooting wireless mesh networks. *SIGCOMM Comput. Commun. Rev.*, 36(5):17–28, 2006.

[18] M. Rodrig, C. Reis, R. Mahajan, D. Wetherall, and J. Zahorjan. Measurement-based characterization of 802.11 in a hotspot setting. In *Proceedings of the 2005 ACM SIGCOMM workshop on Experimental approaches to wireless network design and analysis*, pages 5–10, 2005.

[19] D.-H. Shin and S. Bagchi. Optimal monitoring in multi-channel multi-radio wireless mesh networks. In *MobiHoc*, pages 229–238, 2009.

[20] R. P. Sithamparanathan Kandeepan and I. Chlamtac. Spectrum sensing for cognitive radios with transmission statistics: Considering linear frequency sweeping. *EURASIP Journal on Wireless Communications and Networking*, 2010, 2010.

[21] A. Srinivasan. Distributions on level-sets with applications to approximation algorithms. In *FOCS*, pages 588–597, 2001.

[22] A. P. Streich, M. Frank, D. Basin, and J. M. Buhmann. Multi-assignment clustering for boolean data. In *Proceedings of the 26th Annual International Conference on Machine Learning (ICML)*, pages 969–976, New York, NY, USA, 2009. ACM.

[23] J. Yeo, M. Youssef, and A. Agrawala. A framework for wireless LAN monitoring and its applications. In *WiSe '04:*

Proceedings of the 3rd ACM workshop on Wireless security, pages 70–79, 2004.

[24] J. Yeo, M. Youssef, T. Henderson, and A. Agrawala. An accurate technique for measuring the wireless side of wireless networks. In *the 2005 workshop on Wireless traffic measurements and modeling*, pages 13–18, 2005.