Self Organization of Interfering 802.11 Wireless Access Networks

Bruno Kauffmann, François Baccelli, Augustin Chaintreau, Konstantina Papagiannaki, Christophe Diot

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Self Organization of Interfering 802.11 Wireless Access Networks

Bruno Kauffmann — François Baccelli — Augustin Chaintreau

Konstantina Papagiannaki — Christophe Diot

N° 5649

Août 2005

Thème COM

Rapport de recherche

Self Organization of Interfering 802.11 Wireless Access Networks

Bruno Kauffmann [*] , François Baccelli [†] , Augustin Chaintreau [‡]
Konstantina Papagiannaki [§] , Christophe Diot [¶]

Abstract: The increased popularity of IEEE 802.11 WLANs has led to dense deployments in urban areas. Such high density leads to sub-optimal performance unless the interfering networks learn how to optimally share the spectrum. This paper proposes a set of novel fully distributed algorithms that allow (i) multiple interfering 802.11 WLANs to select their operating frequency in a way that minimizes global interference, and (ii) clients to choose their Access Point so that the bandwidth of all interfering networks is shared optimally. The proposed algorithms rely on Gibbs' sampler and optimize global network performance based on local information. They do not require explicit coordination among the wireless devices. We establish the mathematical properties of the proposed algorithms and study their performance using analytical, event-driven simulations. Our results strongly motivate the need for self-organization strategies in wireless access networks. We discuss implementation requirements and show that significant benefits can be gained even within incremental deployments and in the presence of non-cooperating wireless clients.

Key-words: WIFI, access point, radio channel selection, load balancing, Gibbs' sampler, bandwidth sharing, potential delay fairness.

[*] ENS, (1) bruno.kauffmann@ens.fr

[†] INRIA-ENS, (1), francois.baccelli@ens.fr   (1): 45 rue d'Ulm 75005, Paris, France.

[‡] Intel, (2), Augustin.Chaintreau@intel.com

[§] Intel, (2), dina.papagiannaki@intel.com

[¶] Intel, (2), christophe.diot@intel.com   (2): Intel Research Cambridge, 15 J. J. Thomson Avenue, Cambridge, CB3 0FD, UK.

Auto-Organisation de Réseaux d'Accès 802.11 Interférant

Résumé : Suite à la généralisation de l'usage des réseaux locaux sans fil de type IEEE 802.11, on assiste à l'apparition de zones de haute densité de ces réseaux, notamment en milieu urbain. Une haute densité de points d'accès 802.11 peut conduire à une dégradation significative des performances si les réseaux interférant et leurs clients n'apprennent pas à partager le spectre radio de manière optimale. Dans cet article, nous proposons une famille d'algorithmes distribués qui permettent (i) à des réseaux 802.11 interférant de sélectionner leur fréquence de manière à minimiser les interférences globales, (ii) aux clients de ces réseaux de choisir leur point d'accès de manière à ce que la bande passante des divers réseaux soit partagée de façon optimale. Les algorithmes proposés sont fondés sur l'échantillonneur de Gibbs. Ils permettent une optimisation globale des performances du réseau tout en n'utilisant que des informations locales. En particulier, ils n'exigent pas de coordination explicite entre les points d'accès. Nous établissons diverses propriétés mathématiques de ces algorithms, notamment en ce qui concerne leur convergence, et nous étudions leurs performances de manière analytique et par simulation. Les résultats que nous obtenons montrent clairement l'intérêt de mécanismes d'auto-organisation de tels réseaux d'accès. Nous passons en revue les pré-requis technologiques pour une implémentation pratique de ces algorithmes; nous montrons aussi qu'un déploiement incrémental conduit à des gains en performance significatifs, même si certains points d'accès ou certains clients n'adoptent pas ces algorithmes.

Mots-clés : WIFI, point d'accès, sélection de canaux radio, équilibrage de charge, échantillonneur de Gibbs, partage de bande passante, équité des délais potentiels.

I. INTRODUCTION

Wireless Local Area Networks (WLANs) are currently deployed in a variety of environments such as in university campuses, enterprise buildings, public places, or homes. WLANs are mostly used as a flexible way to access the Internet while allowing for user mobility. Their ease of installation and low cost has led to highly dense IEEE 802.11 deployments in urban areas, where users may be in range of several Access Points (APs) belonging to private or commercial entities.

Coexistence of multiple WiFi devices in an area is bound to lead to high degree of interference and sub-optimal performance due to the limited size of the shared spectrum. Current 802.11 deployments are typically not coordinated and in certain cases not even managed (such as home APs). To make matters worse, current 802.11 equipment typically ships with static configuration that completely ignores the utilization of the radio resources in the area of the AP. It is not uncommon to find areas where multiple APs operate on the same frequency, simply because this was the default factory setting.

In this work we advocate that in order for the WiFi model to survive its success, wireless devices will need to become self-configurable. Consequently, 802.11 devices need to be able to assess the radio environment they operate within and adjust their configuration appropriately. Within this generic framework, we look at ways in which (i) APs can individually identify the optimal frequencies to use, so as to achieve optimal spatial reuse and (ii) clients can identify the "best" APs to affiliate with, such that the resulting configuration of the collection of interfering networks is optimal in terms of bandwidth sharing.

The most popular solution is to use a central location responsible for the optimal configuration of the entire network, operating on information received by every single entity in the network. This approach has two main drawbacks. First, it is not always feasible as all APs do not belong to the same administrative authority. Second, it is not robust to failures and other impairment that are frequent in wireless environments. Instead, we look at fully distributed algorithms that require no coordination between the different network entities (thus avoiding inter-operability issues), and can be deployed incrementally. We are interested in scenarios where each network element (i.e. APs and end users) can reach its optimal configuration simply based on local measurements and minimal information dissemination to/from neighboring devices.

We do not address solutions that require changes at the MAC level. This would not be compatible with incremental deployment on existing 802.11 technologies. Therefore, we focus on solutions that can be implemented through simple software modifications that could be pushed to the wireless devices through firmware upgrades (in line with current proposals within the IEEE 802.11k task group).

We design a set of fully distributed algorithms that rely on Gibbs' sampler to optimize resources in a group of interfering wireless access networks (and their users). We prove the convergence of these algorithms to a global optimum. We propose an implementation through local optimization decisions. Our optimization criterion is the minimal potential delay fairness (defined for wired networks in [12]), which captures the long term rate that a user should expect to receive from a fully saturated network. We show analytically that these algorithms can successfully minimize global interference while optimizing bandwidth sharing across clients. This result stands for both static and dynamic environments, where APs and users may frequently join and leave the network. In case of incremental deployment, our algorithms can have a significant performance benefit even for the first nodes that implement the suggested functionality. Moreover, there is little incentives for people to misbehave, as a more selfish choice does not lead to significant individual performance gain.

The paper is organized as follows. The next section discusses related work. In Section III, we formulate the problem, introduce notation and assumptions. In Section IV, we introduce our distributed channel selection algorithms and establish their optimality analytically in the static case (i.e. when the AP and client populations do not change). The performance of the proposed algorithms is extensively simulated and compared to currently used strategies in Section V. Section VI addresses deployment issues. We demonstrate the flexibility of the proposed solution by relaxing three of the underlying assumptions in the original model in Section VII. We conclude and discuss future work in Section VIII.

II. Related Work

The IEEE 802.11 family of protocols defines a set of frequencies that can be used by wireless devices in their communication. IEEE 802.11b/g operates on the 2.4 GHz spectrum and defines 11 communication channels, 3 out of which are non-overlapping. IEEE 802.11a operates on the 5 GHz spectrum and defines 52 communication channels, 12 out of which are orthogonal. Devices operating on orthogonal frequencies can operate independently. In the opposite case, devices will need to share the wireless resources with other occupants in overlapping frequencies. The way the operating frequency of a wireless device is chosen is, however, typically left to implementation or may be simply statically set by the manufacturer (a typical case for home APs). Users are further associating to APs within range either according to user policies or simply based on the received signal strength, under the assumption that APs with a stronger signal may offer better performance, thus ignoring their load.

Leung and Kim analyze the problem of frequency selection in [11], where they study how to optimally allocate channels to cope with a given offered traffic per AP. They study the complexity of the optimization problem and propose a centralized algorithm to compute channel selection. Similarly a centralized algorithm for the association of clients to APs is formulated by Bejerano et al. in [2] as the solution of an "association control" problem. They analyze its complexity and propose a centralized algorithm based on fractional load balancing and routing, to achieve a near optimal max-min fair allocation of links.

In a similar spirit, numerous papers have addressed optimal channel allocation in wireless meshed networks [8], [9], [15]. The association of users to APs may be seen as a subproblem of the wireless mesh routing problem: each user is allowed to communicate with any AP and needs to choose an optimal link. These algorithms also require a centralized knowledge and computation, that is most of the time not possible to deploy in real life. In [5], Gambiroza et al. propose to achieve fairness in a wireless meshed network by an exchange of control messages, prescribing the rate that should be used on each link of the mesh. In a similar manner, Raniwala [17] proposes that network elements should negotiate channel selection

to maximize the diversity of orthogonal channels used while still preserving the connectivity of the multi-hop architecture. Such protocols for link setup and maintenance are used to build optimal routes in wireless mesh networks. They imply an overhead and a complexity in the wireless cards that may not be needed in the setting that we examine here which is that of interfering WLANs.

To the best of our knowledge, this paper is the first to propose a set of algorithms that simultaneously solves the problems of channel selection and user association in a fully distributed way. Some commercial products[1] claim that they have a solution to these problems. However, the technology used is not disclosed.

The closest problem scenario to our work is described in [1]. The authors look into the problems arising in dense, unmanaged areas, which they call "chaotic". Within this framework, they propose new mechanisms for power and rate control to alleviate the performance degradation inherent in chaotic environments, due to interference. Channel assignment is very briefly addressed in [1], but was found to have limited impact in terms of performance; APs are recommended to select an orthogonal frequency that none of their neighbors uses if possible. Our approach systematically targets channel selection by APs, where channels can be selected among the orthogonal channels or not. Moreover, our work further targets user association which is not addressed in [1].

In terms of our algorithmic contribution, [10] appears to be proposing an algorithm that bears similarities with our work. A method is proposed that can minimize the expected transmission in a single channel based on the estimation of local SINR values. This point is discussed further in this paper, and we compare by simulation in § VI-B.2 our association algorithm and a simplified version of the algorithm in [10].

What makes our approach unique is that we can justify analytically that the algorithms we propose lead to a globally optimal bandwidth sharing where optimality is defined in terms of minimal potential delay fairness [12]; this contrasts the max-min fair sharing of the network's resource, previously proposed. The aforementioned choice is motivated by recent advances in the fundamental fairness effi-

---

[1]see e.g. http://www.propagatenet.com/ .

ciency trade-off in wireless networks (see e.g. [16] and references therein).

III. PROBLEM STATEMENT

As explained earlier, we consider geographical areas where multiple wireless APs are available to users. These APs will most probably interfere, and in many cases, wireless clients will have to choose among multiple wireless networks. Our intent is to provide APs with a completely distributed method to optimally select their operating frequency to minimize global interference, and wireless clients with a completely distributed method to select an AP in order to lead to globally optimal bandwidth sharing. This distributed optimization problem will be addressed in two steps:

- **Maximize the total capacity of the interfering Wireless LANs.** This is achieved by each AP choosing its frequency in such a way that it minimizes interference with neighboring APs.
- **Assign each wireless client a fair share of this capacity.** Clients take into account the amount of interference they experience and generate, as well as the observed *load* of each AP (where the load is defined in a precise way in §III-B) to achieve an optimal bandwidth sharing.

The solution we propose to the problem described above must meet the following challenges. (1) It must be totally distributed in order to allow for self-configuration of wireless APs and clients, in the general case of dynamic populations and non cooperative environments. (2) It should require a minimal information exchange between the network devices. (3) It should lead to an optimal configuration, with respect to a well chosen performance objective, while being robust to the join and leave events of both APs and users. (4) It should be easily deployable on existing equipment, and be able to coexist with non cooperating or mis-behaving equipment. In particular, we assume that any AP may be selected for association and that all antennas are omni-directional. The exact functionality and features required from APs and from clients to support our algorithms and the way for these devices to gather the data needed to run these algorithms will be discussed in detail in §VI.

Note that because of the dynamic nature of the environment, we wish to make the assumption that AP channel selection is independent from the AP's current user population. This channel selection must be in particular independent of the current or average loads of AP. Such a separation is an important feature that allows to have simpler algorithms that remain stable with fast user join and leaves. As a consequence, AP channel selection is decided based upon a global measure of spatial reuse, which is defined precisely in the next section.

Our solution focuses on the optimization of the performance experienced by UDP user traffic on the *downlink* (from the APs to the users). The analysis of uplink traffic and/or of TCP traffic is left for future research.

In the downlink traffic case, we primarily consider the following persistent traffic scenario: users always have some data to download and their rates are only limited by the wireless links they use. In this case, there is always maximal interference between APs. The method that we propose is extended to the case of non persistent traffic in Section VII.

We initially consider our channel selection and host association schemes when the populations of users and APs are fixed. The more realistic case of dynamic populations will be addressed in § V-D.

In the following subsection, we introduce notation and give a formal description of the problem. Then we formalize the bandwidth sharing problem in an 802.11 wireless access network.

A. Problem Formulation

The wireless access network that we consider in this paper is described by :

- A set of APs $\mathcal{A}$.
- A set of users $u \in \mathcal{U}$.
- A set of available channels $c \in \mathcal{C}$.

At first, we assume that these channels do not interfere, this could correspond to channels 1, 6, and 11 of the 802.11b spectrum; the general case (i.e. interfering channels) is amenable to a similar analysis with small modifications as shown in Section VII.

Let $c_a \in \mathcal{C}$ be the channel that is chosen by access points $a \in \mathcal{A}$. We introduce the function $s(a, b)$ that is equal to 1 if a and b are operating on the same channel, and to 0 if they are orthogonal.

Let $a_u \in \mathcal{A}$ be the AP that is chosen by user $u \in \mathcal{U}$. For each AP a, we introduce $\mathcal{U}_a \subseteq \mathcal{U}$ the subset of users associated with a. Note that the collection of these subsets is a partition of the set $\mathcal{U}$. We denote by U_a the number of users associated with a. For all

pairs of users u and v, let $s(u,v)$ be the function whose value is equal to 1 if u and v are associated with the same AP, and to 0 in all other cases.

The AP channel assignment problem consists in choosing a collection $(c_a)_{a \in \mathcal{A}}$ in $\mathcal{C}$ satisfying some given criteria. Similarly we can define the user association problem as the choice of a collection $(a_u)_{u \in \mathcal{U}}$ that verifies certain criteria. These two criteria define a performance objective that we present in Section IV. Before doing so, we describe more precisely our model for APs.

B. Bandwidth Sharing in a Cell

We define a cell as an AP and all the users that are associated to this AP. Within our traffic scenario featuring a saturated AP, it makes sense to assume that bandwidth is shared equally among all users of the same cell as all users achieve the same long term throughput. This is equivalent to assuming that MAC layer mechanisms and higher level protocols achieves a max-min fair allocation inside each cell, as observed for instance experimentally in [4].

Two users associated with the same AP do not in general receive the same signal because of different distance to the AP and varying channel condition. The *auto-rate function* of 802.11 adapts the encoding rate of the transmitter to the channel conditions. It is implemented as a stair function of the SNR. It was shown (see [14]) that this function may be well approximated by a piecewise linear function. For simulations we approximate it as a continuous linear function, with a specific maximum value, as done for example in [16]. However, we note that our algorithms do not rely on a specific relationship between SNR and rate, since they rely on actual measurements.

We assume that information is transmitted to each user in data units of the same length, so that the propagation delay experienced by a data unit sent from the AP to user u is given by:

$$\frac{1}{f(\text{SNR}(u))},$$

where $f(\text{SNR}(u))$ gives the transmission rate on the channel from a to u that is expressed in data unit/seconds.

Within the persistent traffic scenario, the max-min fair allocation of bandwidth in the cell implies that the long term throughput obtained by each user u associated with a is given by:

$$r_u = \frac{1}{\sum_{v \in \mathcal{U}_a} \frac{1}{f(\text{SNR}(v))}} \ . \qquad (1)$$

Note that despite the fact that the time to transmit the same unit of information is different from one user to another in the same cell, all the users of the same cell receive the same long term throughput. The inverse of this quantity will be referred as the *load* of the AP in what follows. This egalitarian feature of 802.11 was recently considered harmful by several researchers, as a unique client with bad channel condition (i.e. low SNR), impacts all users in the same cell [6]. Modifying this mechanism for 802.11 cells was already proposed, to achieve a max-min fair bandwidth sharing based on medium time access, rather than rate. In this paper, we primarily concentrate on the case where 802.11 cells implement a rate-based max-min bandwidth as described above. The case of time-based max-min sharing can however be handled in the same way as shown in Section VII.

Within our persistent traffic scenario, the SNR received by user u on the downlink is given by:

$$\text{SNR}(u) = \frac{P_a(u)}{N_u + \sum_{b \in \mathcal{A} \,|\, b \neq a} s(b,a) P_b(u)} \ , \qquad (2)$$

where we denote by $P_a(u)$ the power received by user u from an AP a, which typically depends on the transmission power of a and its distance to u; N_u denotes the thermal noise seen by client u.

Note that this previous expression for the SNR experienced by user u relies on an approximation: the interference from another cell is always estimated by the interfering signal created by its associated AP. This neglects periods of up-link traffic, where packets are transmitted by users.

IV. OPTIMAL BANDWIDTH SHARING USING A GIBBS SAMPLER

In this section we introduce the two algorithms that we propose to achieve the two objectives described earlier, i.e. the maximal spatial reuse among channels selected by APs, and fair share of the wireless bandwidth among all wireless clients in the network. Both are modeled as the minimization of a global energy function that is defined on state variables. The state variable of an AP is its channel, and the state variable of a user is its associated AP.

We present a series of distributed algorithms, based on Gibbs sampler, that achieve this optimization. A Gibbs sampler is a mechanism implemented by each AP/user to update its state. It mimics locally the evolution of a Markov Random Field. It is possible to control the global evolution of this Random Field with a temperature parameter T. This can lead, under some condition, to a state realization in APs/users that achieves a global minimal energy.

A. Performance Objectives

We describe in this section the detailed performance objectives that are met by our scheme. We start with the definition of a maximal spatial reuse achieved by our channel selection. We then introduce a fairness criterion that is used to define the optimal bandwidth sharing that is implemented by our association algorithm.

1) AP channel selection: The interference experienced by an AP a on channel c_a is defined as the amount of power it receives from other APs operating in the same frequency, plus the inherent noise on channel c_a as perceived by the AP a (which may include for instance transmissions of non-802.11 devices). Consequently, when it comes to the selection of a channel $(c_a)_{a \in \mathcal{A}}$ by an AP a, our objective is to minimize the following energy function:

$$\mathcal{F}\left((c_a)_{a \in \mathcal{A}}\right) = \sum_{a \in \mathcal{A}} \left(N_a + \sum_{b \in \mathcal{A} \,|\, b \neq a} s(a,b) P_b(a) \right), \quad (3)$$

where N_a denotes the level of thermal noise estimated locally by this AP on its channel, and $P_b(a)$ denotes the power of the signal received in a from AP b. In the rest of this paper we do not write $(c_a)_a$ as arguments of $\mathcal{F}$ in order to simply the notation. However, it is important to keep in mind that $\mathcal{F}$ always depends on these variables. $\mathcal{F}$ can be further expanded as:

$$\begin{aligned} \mathcal{F} &= \sum_{a \in \mathcal{A}} \left(N_a + \sum_{b \in \mathcal{A} \,|\, b \neq a} s(a,b) P_b(a) \right) \\ &= \sum_{a \in \mathcal{A}} N_a + \sum_{\{a,b\} \subseteq \mathcal{A}} s(a,b)(P_b(a) + P_a(b)) \\ &= \sum_{\mathcal{B} \subseteq \mathcal{A}} V(\mathcal{B}), \end{aligned}$$

where the function V is defined for all subsets of $\mathcal{A}$ by

$$\begin{aligned} V(\mathcal{B}) &= N_a & \text{for } \mathcal{B} = \{a\}, \\ V(\mathcal{B}) &= s(a,b)(P_a(b) + P_b(a)) & \text{for } \mathcal{B} = \{a,b\}, \\ V(\mathcal{B}) &= 0 & \text{for } |\mathcal{B}| \geq 3. \end{aligned}$$

Following the terminology of Gibbs field theory (see [3] Chapter 6), this rewriting shows that the energy function $\mathcal{F}$ derives from the potential $\{V(\mathcal{B})\}_{\mathcal{B} \subseteq \mathcal{A}}$. In particular, we can define the local energy of an AP a as:

$$\mathcal{F}_a = \sum_{a \in \mathcal{B}} V(\mathcal{B}) = N_a + \sum_{b \neq a} s(a,b)(P_b(a) + P_a(b)).$$

Assuming that each AP uses the same nominal power to transmit, we have

$$\mathcal{F}_a = N_a + \sum_{b \neq a} 2s(a,b) P_b(a).$$

2) User association: We assume that a user would always try to associate with the AP that offers the best long-term throughput. Massoulié and Roberts [12] showed that an interesting optimization criterion for such an allocation is one that achieves *minimal potential delay* fairness, which can be expressed as an energy function:

$$\mathcal{E}\left((a_u)_{u \in \mathcal{U}}\right) = \sum_{u \in \mathcal{U}} \frac{1}{r_u}. \quad (4)$$

The potential delay of a user is defined as the inverse of its rate and may be interpreted as the delay for the network to transmit one unit of information for this user. Hence, minimizing this sum is equivalent to minimizing the sum of delays for a unit of information to be transmitted for each user. This property motivates the name of this fairness criterion; it also provides an intuitive interpretation of the energy as a cost function of the network.

We denote by $\{u,v\} \subseteq \mathcal{U}$ a pair of distinct elements of $\mathcal{U}$. The energy function $\mathcal{E}$ can be expanded as [2]:

$$\begin{aligned} \mathcal{E} &= \sum_{u \in \mathcal{U}} \frac{1}{r_u} = \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{U}} \frac{s(u,v)}{f(\mathrm{SNR}(v))} \\ &= \sum_{u \in \mathcal{U}} \frac{1}{f(\mathrm{SNR}(u))} + \sum_{u,v \in \mathcal{U} \,|\, u \neq v} \frac{s(u,v)}{f(\mathrm{SNR}(v))} \\ &= \sum_{u \in \mathcal{U}} \frac{1}{f(\mathrm{SNR}(u))} + \sum_{\{u,v\} \subseteq \mathcal{U}} \left(\frac{s(u,v)}{f(\mathrm{SNR}(v))} + \frac{s(u,v)}{f(\mathrm{SNR}(u))} \right). \end{aligned}$$

---

[2] As for the previous algorithm, we do not explicitly write $(a_u)_u$ as arguments of this function to keep notation light, but they are always implicitly included in the following mathematical expressions.

This energy can be shown to derive from a potential function V defined on the subsets of U:

$$\mathcal{E} = \sum_{\mathcal{V} \subseteq \mathcal{U}} V(\mathcal{V}), \qquad (5)$$

where V is defined for every subset $\mathcal{V}$ of $\mathcal{U}$ as

$$
\begin{aligned}
V(\mathcal{V}) &= \frac{1}{f(\text{SNR}(u))} & \text{for } \mathcal{V} = \{u\} \\
V(\mathcal{V}) &= \frac{s(u,v)}{f(\text{SNR}(u))} + \frac{s(u,v)}{f(\text{SNR}(v))} & \text{for } \mathcal{V} = \{u,v\} \\
V(\mathcal{V}) &= 0 & \text{for } |\mathcal{V}| \geq 3 .
\end{aligned}
$$

We can define the local energy of a given user as:

$$\mathcal{E}_u = \sum_{u \in \mathcal{V}} V(\mathcal{V}) = \sum_{v \in \mathcal{U}_a} \frac{1}{f(\text{SNR}(v))} + U_a \frac{1}{f(\text{SNR}(u))} . \qquad (6)$$

Let us quickly comment on these expressions. The energy $\mathcal{E}$ is a sum of local energies defined on subsets of users. Indeed it is positive only for subsets containing only one user, or two users associated with the same AP. The local energy of user u is the sum of the local energy for all subsets that contain u. This can be interpreted as the part of the global energy for which u can have an impact. It contains in particular the inverse of its rate (the left hand term of the RHS of (6)), as it was clear in the first expression of the energy function (see (4)). In addition, it contains another term that represents its *social cost*, which translates the fact that traffic toward this user increases the potential delay of all other users of this cell. This social cost is that on the right of the RHS of (6).

B. Algorithms for AP Channel Selection

The general idea is to trigger AP channel transitions according to a rule that drives the network to a state that minimizes energy (and thus converges to the optimal channel selection). The difficult problem is to orchestrate transitions from different APs that are all sharing the same spectrum (at different levels of interaction). In our solution each AP maintains the value of an exponential timer, with mean t_a; whenever this timer expire, it follows the following transition:

Algorithm 1 *(AP transition)*

1) *Compute the temperature parameter:* $T = \frac{K}{\log(2+t)}$.

2) *For all channels c, compute the local energy experienced by a on this channel :*

$$\mathcal{F}_a(c) = N_a + 2 \sum_{b \in \mathcal{A} \,;\, c_b = c} P_b(a) .$$

3) *For all channels c, compute the probability*

$$\pi(c) = \left(e^{-\frac{\mathcal{F}_a(c)}{T}} \right) \Big/ \left(\sum_{c \in \mathcal{C}} e^{-\frac{\mathcal{F}_a(c)}{T}} \right) .$$

4) *Sample a random variable with law π and choose a channel according to this random variable.*

Here, t is an *age variable* (that roughly represents the time elapsed since the most recent reconfiguration of the network) which is re-initialized to zero by certain events to be described later. K is a constant that is chosen for initial condition.

The rationale for this algorithm is the following. The channel chosen by each AP is a state variable. Therefore, we have a collection of variables distributed on the plane. We denote this collection by $c = (c_a)_{a \in \mathcal{A}} \in \mathcal{C}^{\mathcal{A}}$. For a given temperature T, and an energy function $\mathcal{F}$, the distribution π defined by

$$\pi(c) = e^{-\frac{\mathcal{F}(c)}{T}} \Big/ \left(\sum_{c' \mathcal{C}^{\mathcal{A}}} e^{-\frac{\mathcal{F}(c')}{T}} \right)$$

is called the Gibbs distribution associated with this energy. It is proven in §IV-D that for a fixed topology this algorithm combined with the logarithmic cooling performed in the first step allows the state variables to converge to the state of minimal energy.

This algorithm can be adapted to a slowly varying dynamic topology. To deal with the case of a dynamic population of APs, each AP should maintain the list of the APs that it receives with their respective strength. If this list changes (because of an AP joining or leaving), the age variable of the AP is re-initialized to zero. For highly dynamic topologies a reasonable and practical solution, used in the rest of this article, is to fix the temperature parameter to a given well chosen constant.

A greedy algorithm : The previous probabilistic algorithm is rather complex as it needs to maintain the value of t. In practice we could also use the following simpler greedy local optimization.

Algorithm 2 *(deterministic AP transition)*

Choose channel c as $c = \text{argmin}_{c \in \mathcal{C}}(\mathcal{F}_a(c))$

Algorithm 2 may be thought of as Algorithm 1, when the temperature is fixed to 0 in the first step. It is a very good approximation of Algorithm 1 if t is large. Algorithm 2 may then be interpreted as a speedup version of Algorithm 1: instead of a logarithmic temperature decrease, which drives the distribution towards minimal energy states, this algorithm chooses the state of minimum local energy for each transition.

The difference between the performance of the two algorithms is that the first can be shown to converge eventually to a state of *minimal interference*, for any fixed topology. The second algorithm however can get blocked in local minima of the energy $\mathcal{F}$ (see §IV-D). Choosing to implement the second one is therefore trading long term efficiency for quick improvement and simpler implementation. Extensive simulations have shown that in all the cases that we studied, these local minima provide excellent approximations of the optimum obtained by the first scheme.

The extension of this deterministic algorithm to the case of a dynamic population of APs can be done as before. Assuming that each AP maintains a list of its closest neighbors, it might decide to trigger a transition in case it senses a change. This implementation is simpler as the value of t does not need to be maintained nor re-initialized in each AP.

C. Algorithms for User Association

The algorithm for user association is very similar to the one proposed in the previous section. This algorithm assumes that each AP has selected its channel. After collecting information from APs (number of users and load) on the different channels, each user can compute the local energy it would experience when associated with AP a for all a within range. Each user maintains an age variable t and an exponential timer with mean t_u. Whenever its timer expires, each user follows.

Algorithm 3 *(User Transition)*
Follow the same steps as in Algorithm 1 and choose AP a to associate with probability π defined from energy $\mathcal{E}_u(a)$.

A greedy version of this algorithm may be similarly defined as follows:

Algorithm 4 *(User deterministic transition)*
Choose AP a as

$$a = \text{argmin}_{a \in \mathcal{A}} \left(U_a \frac{1}{f(\text{SNR}(u))} + \sum_{v \in \mathcal{U}_a} \frac{1}{f(\text{SNR}(v))} \right) .$$

On a fixed topology, Algorithm 3 may be proved to eventually converge to the association that achieves the minimal potential delay fairness. Its simplified version, Algorithm 4, could be blocked in local minima of the energy $\mathcal{E}$, but simulations show that it performs extremely well in practice. Note that the population of users may vary quite rapidly compared to the set of the active APs. Algorithm 4, which is simpler and more reactive, seems to be a good choice for robust quasi optimality.

D. Analysis for Static Population

Theorem 1 *For a fixed populations of APs and users that implement Algorithm 1 and 3, there exists a value of K such that the AP channel selection and client association verify:*

$$\mathcal{F}\left((c_a)_{a \in \mathcal{A}}\right) \to \min_{(c_a)_{a \in \mathcal{A}}} \mathcal{F}$$
$$\text{and } \mathcal{E}\left((a_u)_{u \in \mathcal{U}}\right) \to \min_{(a_u)_{u \in \mathcal{U}}} \mathcal{E}$$

as time goes to infinity.

Proof: Algorithm 1 and 3 use a Gibbs sampler together with a decrease of the temperature parameter. This is a typical simulated annealing procedure. From Theorem 8.1 in [3], it follows that for a logarithmic decrease of T, the channel allocation has a limit distribution that is only putting positive probability on states of global minimum energy. ∎

Theorem 2 *We now consider fixed populations of APs and clients which implement Algorithms 2 and 4. The resulting AP channel selection and client association are such that*

$$(c_a)_{a \in \mathcal{A}} \to (\tilde{c}_a)_{a \in \mathcal{A}}$$

$$\text{and } (a_u)_{u \in \mathcal{U}} \to (\tilde{a}_u)_{u \in \mathcal{U}}$$

as time goes to infinity, where $\tilde{c}$ and $\tilde{a}$ are local minima, in the following sense:

For any $a \in \mathcal{A}$, $\tilde{c}_a = \text{argmin}_{c \in \mathcal{C}} \mathcal{F}_a(c)$

For any $u \in \mathcal{U}$, $\tilde{a}_u = \text{argmin}_{a \in \mathcal{A}} \mathcal{E}_u(a)$.

Proof: The proof of this theorem is much simpler. We include it here since it nicely illustrates the combined behavior of the two algorithms. We consider first the sequence of channel allocation of APs. From the very definition of Algorithm 2, the global energy function $\mathcal{F}$ can only decrease after each transition, since the current AP transition is bound to lead to a decrease in the sum of energies of all subsets containing that AP, and that the energy of other subsets is not changed. As the state space of AP channels is finite, it is certain that this sequence is constant after a finite number of steps, in a state where no single transition of an AP can decrease the local energy function. Once the sequence of channel allocation is constant, the same argument can be used to prove that the sequence of user associations is also necessarily constant. ∎

V. Empirical Evaluation

For the evaluation of the proposed algorithm, we built our own event-driven simulator using C. Given a network topology of APs and wireless users, as well as their power specifications and the number of orthogonal channels in the spectrum, the simulator implements the proposed algorithms by constantly updating the state of the overall network. Each device (AP or wireless user) in the network is equipped with an exponential timer which triggers the transition as defined by Algorithms 1-4. On all simulations presented an AP makes a transition on average every 3 hours, while users are expected to make a transition on average every 15 minutes. The outputs of the simulator are a map of the channels used by APs and their associated wireless users in time, accompanied by the current value of the energy functions, and the long term rate that each wireless user can achieve in the given topology.

As opposed to packet level simulation tools, such as the ns-2 event driven simulator, our methodology ignores packet level effects. At the end of this section we compare our Gibbs simulator with ns-2 under a simple scenario that uses a small population of APs and users to examine the impact of the simplifying assumptions made. Given the memory requirements of ns-2 when large topologies are taken into account, any such scenario is going to be investigated using the Gibbs simulator alone.

A. Methodological Approach

The outcome of the proposed algorithms is likely to be affected by several parameters which will be investigated in this section. Firstly, the topology of the nodes can have a significant impact since our algorithm is designed to resolve conflicts in areas of high radio interference, i.e where nodes are clustered. To address this aspect of the problem, we look at topologies where APs and wireless users are distributed in a square according to two different distributions:

- **Homogeneous topology**: the locations of APs and wireless users are sampled according to independent Poisson point processes (PPP) in the square. The intensities of these PPPs are chosen to be such that the square in question has 500 APs and 5000 wireless users on average. Note that PPPs are bound to lead to certain areas of higher concentration.
- **Sporadic topology**: same as above but now users are sampled according to a non-homogeneous PPP that results in regions with very high density of users. We configured 10% of the APs on the plane to have 10 times higher user intensity than the global intensity.

For ease of comparison, the same overall mean numbers are used in all cases: 500 APs distributed in a square containing 5000 users.

A second aspect of the problem that may lead to differences in terms of performance has to do with how dynamic the overall topology is in time. We are going to look into two different cases, that of (i) a static topology, and (ii) a dynamic topology, where users and APs may join and leave the network across time. To facilitate the comparison between the two aforementioned cases we implement changes such that the overall number of APs and users remains the same. More specifically, every "leave" event by an AP or user is followed by an equivalent "join" event (somewhere on the plane) so that the population intensities are kept constant over time. Deletions and additions are made at random. The frequency of changes in the topology is expressed using the percentage of the AP population (respectively user population) that changes between two AP transitions (respectively user transitions) and is varied across our experiments. The case where the overall number of network devices is changing with time was also studied by simulation, but it is more difficult to

interpret with regard to our performance objective given the changing population.

In order to demonstrate the benefits of the proposed algorithms we compare its performance with what could be considered the state of the art channel assignment and user association algorithm. We assume that APs select their channels randomly (an assumption that may lead to a much better baseline than what could be found in real wireless environments due to the factory default frequencies), and users affiliate with the AP with the strongest signal strength.

B. Simulator Configuration

The formulation of the energy functions $\mathcal{F}$ and $\mathcal{E}$ depends on estimates of the power received at a from AP b, $P_b(a)$, the noise level N_a, and the relationship of SNR with distance. Our Gibbs simulator uses the 2-ray ground model for propagation, and is calibrated against ns-2. More specifically, we have tuned power, and noise levels such that a single user associated with an AP at a particular distance receives the same throughput in our Gibbs simulator as he/she would in ns-2 (we assume that the rate of a user is proportional to its SNR, and set the scaling parameter accordingly). The SNR of a user u is then given by

$$\text{SNR}(u) = \frac{P|a-u|^{\beta}}{N_v + \sum_{b \in \mathcal{A} \,|\, b \neq a} s(a,b)P|b-u|^{\beta}}, \ \beta = 4. \tag{7}$$

We should note at this point, that the Gibbs simulator allows further flexibility to look into the effect of specific simplifying assumptions, such as the one of equal power transmitted by each AP and user. In fact, multipath effects can be easily integrated as follows:

$$\text{SNR}(u) = \frac{PF_{a,u}|a-u|^{\beta}}{N_v + \sum_{b \in \mathcal{A} \,|\, b \neq a} s(a,b)PF_{b,u}|b-u|^{\beta}}, \tag{8}$$

where $F_{a,u}$ can be independent and identically distributed random variables, emulating the effect that multipath shadowing may have on the transmission range of APs and users.

C. Static Network Topology

In this section we discuss the performance of the proposed algorithms when the network topology remains static.

1) AP Channel Selection: If all APs are assumed to initially use a randomly selected channel, out of three orthogonal channels (resembling the IEEE 802.11b environment), we find that the Gibbs simulator is leading eventually to a 20% reduction in terms of global energy in the channel selection step. The effect of this algorithm alone for a static network with sporadic topology is represented among others in Figure 1 (left).

We observe that APs converge to the optimal frequency after 6 hours, that is after 2 transitions on average. The APs receive the majority of the gain through a single transition. This statement holds true for both sporadic and homogeneous topologies (not presented here). As will be seen in the next section, such a gain may be considered moderate when compared to the actual benefit gained through intelligent user association. This result confirms previous observation such as the one made in [1].

2) Wireless User Association: The second step of our self-organization mechanism relies on users intelligently affiliating with APs in range to achieve optimal load sharing. Such an association can be done either based on received signal stength (e.g. state of the art), or using Algorithm 4. We present the outcome of our user association algorithm for a sporadic topology in terms of average potential delay per user in Fig. 1 (left), with and without the APs channel selection algorithm. This algorithm exhibits a bigger improvement: more than 40% reduction of the average potential delay per user, when only Algorithm 4 is used, and more than 50% when it is used in conjunction with Algorithm 2. This improvement comes also faster, more than 30% reduction observed after a couple of user transitions.

The association maps for different algorithms are presented in Fig. 2 (the AP channel assignment on this map is the result of Algorithm 2) Each of these figures shows the locations of the APs, and their associated users. The two graphs are color-coded such that darker regions on the plane correspond to areas where users receive lower throughput. If clients associate with their closest APs (based on strongest received signal), then depending on the distribution of the users and APs on the plane, there may exist regions of dense user population, and therefore low throughput. Alg. 4 is able to spread the load across the entire network. Some wireless users belonging to a region with a dense population now associate with more remote APs that may offer better performance
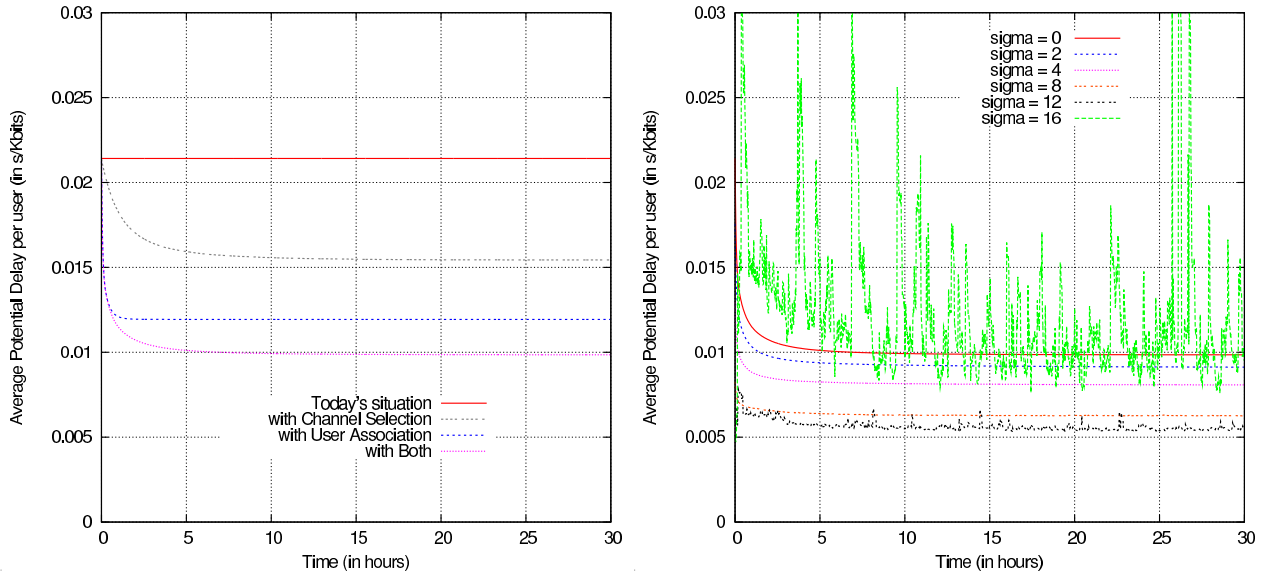
Fig. 1.   Average Potential Delay per user seen as a function of time for a static topology: comparison of different algorithms (left), effect of shadowing (right).

despite their distance. The natural consequence is a more balanced distribution of throughput in this region.

Figure 3 presents the distribution of throughput for the scenarios in Fig. 2. Each plot gives the empirical distribution of throughput obtained by the entire user population in log-log scale. We observe that closest association leads to a highly dispersed throughput distribution. Throughput values range across two orders of magnitude (from 10 Kbps to 1 Mbps). Given that we are looking at a sporadic topology, closest association can lead to a significant number of users that receive low throughput due to the high user concentration around an overloaded AP. In contrast, Alg. 4 is capable of leading to a more even distribution of throughput, eliminating its lower end. However, we note that Alg. 4 also eliminates the very high throughput users, since their APs now need to serve other users previously associated with overloaded APs.

Note that the mean overall throughput is smaller if our scheme is implemented. This should not come as a surprise, as for closest association, each user is associated to the AP with the highest signal strength, such that the total capacity of the network, taken as the sum of the rate achieved, is the highest. Such a maximization leads to gross unfairness, as seen here, and as it was shown several time before (see e.g. [16]), and it is not suitable as a performance objective for WLANs.

We repeated similar simulations on an homogeneous topology. Alg. 4 is still able to lead to a tighter throughput distribution with a shape similar to that of Fig. 3(b), thus improving the performance that users at the low end would experience.

3) Shadowing: To study the effect of multipath loss exponent on links between APs and users, we consider SNR given by (8), where F follows a log-normal distribution with parameters (ϵ, σ^2), meaning that the random variables with law F are of the form $X = 10^{(\epsilon+\sigma\xi)/10}$, where ξ is a standard normal random variable. Note that ϵ, σ^2 can be interpreted as mean and variance of X expressed in dB. In the simulations, we choose $\epsilon = 0$ and let σ vary.

The results are presented in Fig. 1 (right) for σ up to $16dB$. Note that, from the perspective of the same user, the effect of multiple paths on the power of signal received from several APs is different. When σ increases, this effect tends to differ significantly among APs due to the random variable, especially when this variable is chosen independently among APs, as done here. One consequence is that in some cases the SNR received by a user from its optimal AP may be improved. This is observed for instance on the performance of the association using only signal strength, that corresponds to the initial state and is shown in Fig.1 (right) at the time t=0.

We see that our algorithms offer to less and less improvement, compared with closest association, when σ increases. This is caused by the loss of locality in the optimization problem, as the value of
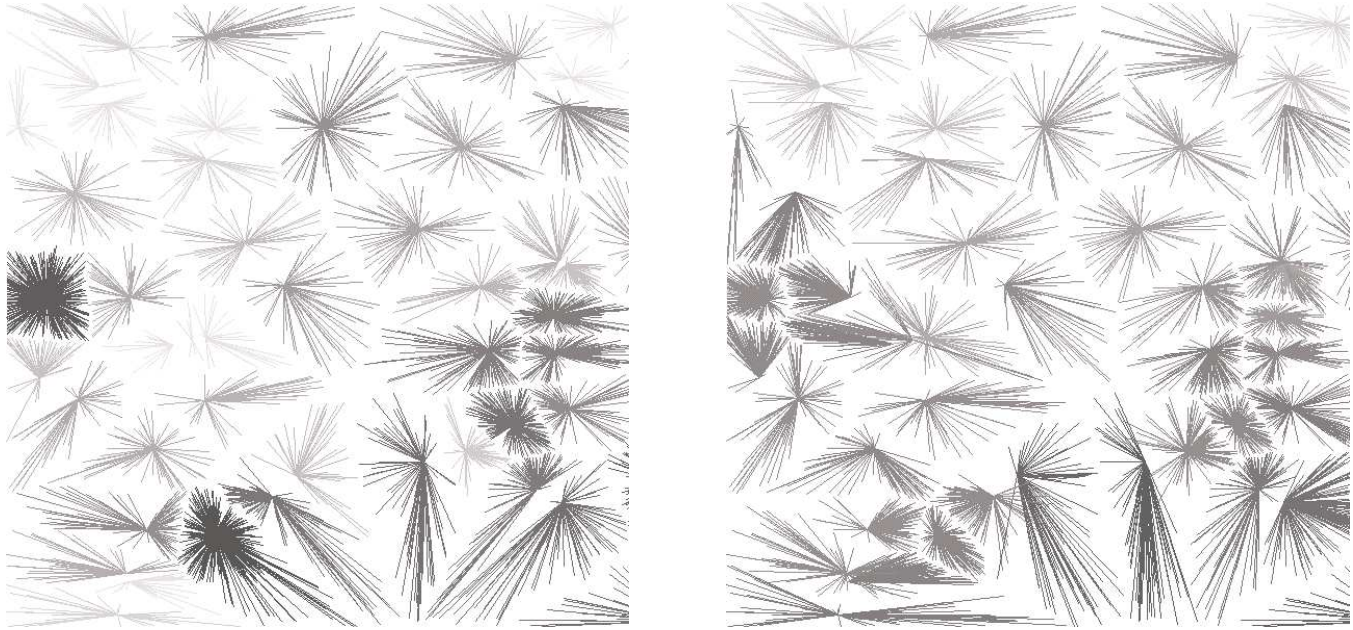
Fig. 2.   Maps representing the wireless users and the links that they select; the color represents the rate obtain by users (darker lines represent links with worse performance): (a) Association to the closest AP (left); (b) Association via Algorithm 4 (right).

the random variable F becomes more dominant in the value of the power received, as opposed to the attenuation function given by the distance. On the other hand, our algorithms still perform reasonably well in improving performance, for moderate value of σ (up to 12dB).

D. Dynamic Topology

The operation of the proposed algorithms relies on exponential timers on each AP and user. Whenever a timer expires, the network device assesses the "state" of the different channels and decides whether to make a transition or not. In the previous section we showed that in the case of a static topology, the algorithm converges quickly to a globally optimal state. The question we address here is whether this statement remains true when the topology itself changes with time.

In the simulation of the dynamic topology case, we use exponential timers to alter the topology. An AP topological change occurs on average every 15 min and a user topological change every 90s. The time between topological changes for the AP (resp. user) population is hence on average ten times smaller than the time between two AP (resp. user) transitions. Each of these topological change events may change between 1% and 7% of the populations of APs (resp. users), so that the proportion of the

population that changed between two transitions is between 10% and 70%.

The join and leave events have the following effect: when some user joins, it automatically triggers a transition to choose an AP association. A user leaving has no other effect than improving the throughput of the other users in the same cell. When an AP joins, it triggers an immediate transition to choose its channel, but users may not immediately associate with it. When an AP leaves, all its users immediately trigger a transition in order to choose another AP to associate with.

Results are shown in Figure 4 for a sporadic topology (results for a homogeneous topology are very similar). From the plot, we can see that the energy is still very much improved and that it stays within 10 % of the optimal allocation for a variety of levels of disruption. This is an encouraging result as it means that the optimality of the network state is not severely impacted by join and leave events. As expected, AP join and leave events have a much stronger disrupting effect on the energy and hence on the rates achieved by the population of users; however, these events occur on a larger time-scale than user join and leave, and they may be taken care of via a more elaborate transition mechanism allowing the network to adapt more quickly when such events occur. Even in highly dynamic cases, our algorithms still significantly improve fairness, compared with
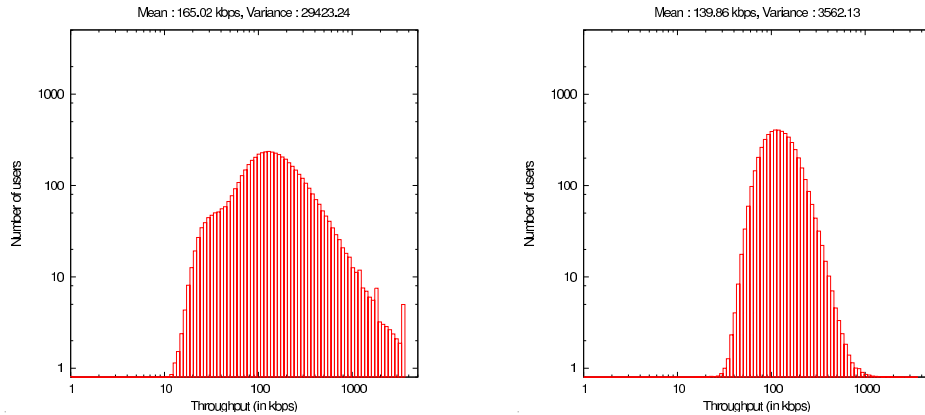
Fig. 3. Histograms of the distribution of long-term throughput, in the case of a sporadic topology: (a) closest association (left); (b) association using Algorithm 4 (right). AP channel selection according to Algorithm 2.

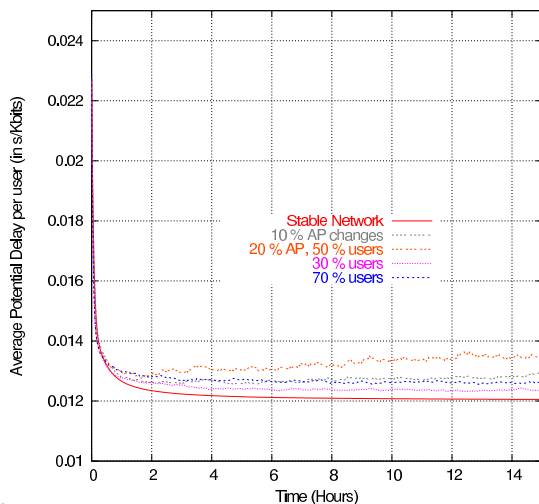the initial default configuration where users associate with the closest AP.

Fig. 4. Average potential delay per user as a function of time for a dynamic topology with different amounts of APs and user changes between two transitions.

E. Comparison with Packet-Level Simulations

All previous results are obtained using the Gibbs simulator. As mentioned before, our simulator does not take into account any packet level effects or MAC layer specificities. In order to identify whether such simplifying assumptions impact our findings, we compare the quantitative results obtained using the Gibbs simulator against the performance that would be obtained with ns-2.

More specifically, we run a small scale simulation using the Gibbs sampler (50 APs and 500 users) and obtain some final snapshot of channel assignments and user associations. We then input
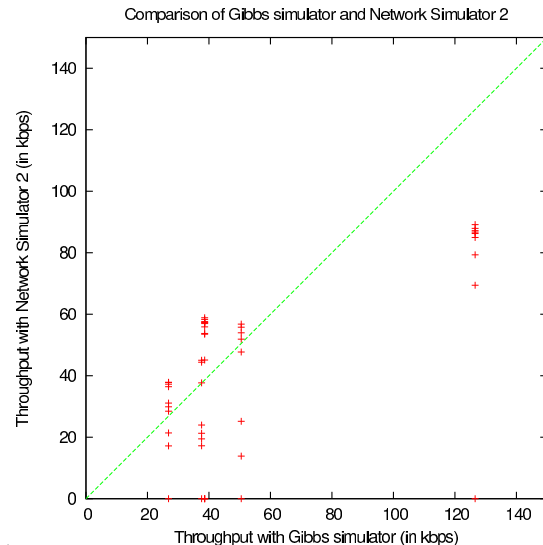

Fig. 5. Comparison of throughput achieves by 50 users in a 5 AP network using the Gibbs simulator and ns-2.

the same topology in ns-2[3]. We initiate exponential traffic sources from APs to their users using 1470 Byte packets (burst and idle times are set to 100ms). The APs' queue length is set to 100 packets and simulations are run for 300 seconds. We compare the throughput value achieved by each user through ns-2 with the respective values obtained using the Gibbs simulator in Fig. 5. Despite the simplicity of the Gibbs simulator, the results are highly accurate; this accuracy comes with the additional benefit of increased scalability and speed.

---

[3]In the ns-2 simulations, we use the default Wireless 802.11b MAC with RTS/CTS and a fixed transmission rate of 11 Mbps. To simulate interfering APs, we use the Uniform Error Model on both AP incoming and outgoing packets. Lastly, to model propagation, we used the Shadowing model with a path loss exponent $\beta = 4$, and a shadowing standard deviation $\sigma(dB)$ =12.

VI. PRACTICAL REQUIREMENTS

The aim of the solution proposed in this paper is to alleviate the performance degradation experienced in today's dense 802.11 environments due to increased interference. Therefore, our two major requirements are: i) the scheme should be implementable in today's technology requiring simple firmware modifications, and ii) it should be incrementally deployable. In this section we will look into these two requirements in detail, describing the implications that our scheme could have on wireless card functionality and the potential benefits gained through the gradual adoption of the proposed scheme.

A. Technological Requirements

The proposed algorithm relies on the following basic principles:

1) APs and clients run exponential timers at the expiration of which they evaluate whether a transition is needed.
2) All wireless devices are able to scan all channels to collect information on all possible "energy states". Such a task can be performed using a second radio or not.
3) For the APs, there is an "energy state" for each channel, that can be evaluated using the power received by all other APs within range, for that particular channel.
4) For a user, there is a different "energy state" for each channel and AP within range on that channel. Each of these states depends on the SNR received from the AP within range operating on this channel, the number of users on that AP and its associated load.
5) Each AP is able to notify all its users about an upcoming change in its operating channel.

Except from the second item, which is already supported in today's 802.11 equipment, all other functionality could be part of software upgrades to the WLAN devices. Some of these changes, moreover, have already been considered for standardization within the IEEE 802.11h and 802.11k task groups. The former defines the mechanisms that need to be implemented by an AP for Dynamic Frequency Selection (DFS) and Transmit Power Control (TPC). The driving need behind 802.11h is the safe coexistence between radar and WLAN devices if they happen to use the same frequency. Within the

proposed standard APs can initiate channel switch announcements to their users so that they vacate a frequency that is used for radar communication. A similar principle could be employed by our scheme in case a change in frequency is deemed desirable by the AP. The IEEE 802.11k further defines a framework to facilitate radio resource management within which WLAN devices exchange statistics, say to make more informed roaming decisions. The number of users supported by an AP is a part of the 802.11k specification and the long-term delay/load could easily be added and communicated to the hosts inside the beacon probes. Notice that 802.11k is meant to describe functionality that can be implemented in software so that changes can be quickly disseminated to currently functional WLAN devices.

Consequently, we believe that the proposed algorithm could be easily integrated in the framework already considered by the standardization bodies. In such a case, a natural second issue to discuss concerns the feasibility of incremental deployment, and the potential danger arising from misbehaving wireless devices.

B. Unresponsive or Selfish APs and Users

The friendliest environment for the deployment of our algorithms with obvious benefits both to the network operator and the users is the one of a single administrative authority, e.g. campus or enterprise network. In such a case, all WLAN devices are likely to be upgraded at the same time establishing the information flow needed by our scheme. However, the major benefits from a distributed scheme like the one presented here arise in environments where the WLAN devices are not managed by the same entity, such as hotspots, neighborhood networks, etc. In this latter case, APs and users experience interference from other 802.11 devices that fall under different administrative authorities, and may not support the functionality required by the algorithms, either because they have not been upgraded or because they choose not to.

Within a non-cooperative environment APs may decide not to adjust to the overall network state and remain on the same channel, as done today. Furthermore, users may choose to associate to specific APs in the set of those within range, due to security considerations or charging schemes employed by the APs. These two aforementioned issues fall

outside the scope of this work since they depend more on policy rather than technological aspects of the problem. From a technical standpoint, however, interesting non-cooperative cases arise when a user either does not employ the suggested functionality, which we call unresponsive user, or prefers to act selfishly and to always affiliate with the AP that offers the lowest potential delay, and therefore the highest long-term throughput.

If cooperating nodes/APs in an incremental deployment end up receiving worse performance when complying to the scheme, there is no incentive to adopt the described mechanism. Moreover, if acting selfishly in such an environment leads to significant gains, then counter mechanisms may need to be developed. Given that the benefit from our distributed channel assignment algorithm are self-evident this section focuses on the users.
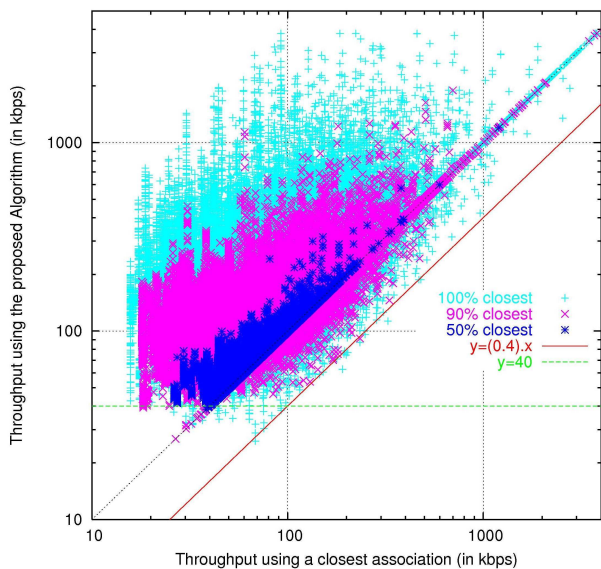

Fig. 6. Comparison of throughput obtained by a user for closest association and association with Algorithm 4, in the case of several levels of deployment

1) Unresponsive Users: Focusing on the same AP/user topology, Figure 6 compares the throughput obtained by a user if he/she implements today's default closest association mechanism, and if he/she decides to switch to Algorithm 4. Each point represents the two possible throughput values achieved by a given user under the two difference scenarios. We further include three sets of points that correspond to different numbers of users already using Algorithm 4. The largest set of points in this figure, in light color, represents cases when a single user adopts our algorithm in an otherwise

legacy environment. Most of the users (80% in our simulations) gain significant performance improvement; when switching to our algorithm, users always avoid the worst-case throughput values. In some less likely cases (6% of the studied cases), it is however possible for users to lose throughput by applying the algorithm, going typically from a very large value to a smaller one. However, this decrease is moderate and amounts to a 60% decrease in the most extreme cases.

In Figure 6, the set of points in the middle, represented in grey, shows the throughput achieved by a new user switching to our scheme when 10% of the overall user population employs Algorithm 4. In comparison to the first case, gains are more moderate for a smaller number of cases (66%), while 8% of the cases may lead to a smaller throughput value, even though such a loss is typically smaller in magnitude than in the case where all users implement the legacy association algorithm.

In the same Figure 6 we also study the benefit gained by a new user when 50% of the population already uses Algorithm 4 for their association. As shown by the set of black points, gains and losses under such conditions are more moderate. In essence, at this point the system has reached a state of optimal load balancing across space, and users that adopt Algorithm 4 are bound to receive their fair share.

Indeed for this last case, a mixed population where only half of the clients use our association algorithm, the histogram of the obtained throughput, shown in Figure 7, is already matching quite well the histogram of the throughput obtained for a population with all users associating with our algorithm. The histogram for users that did not choose to adopt our algorithm is in fact very similar. It appears that only half of the population implementing our scheme is sufficient to achieve a reasonably accurate fair resource allocation. This proves that optimal bandwidth sharing may be reasonably attained with proportion of user that use legacy mechanism.

2) Selfish Users: In the previous section we considered a mixed user population where some users associate to the closest AP, today's default option, while other implement Alg. 4. In almost all cases, users have a strong incentive to adopt our algorithm as it improves their performance. However, one could imagine that some wireless users might be interested in maximizing their performance but not
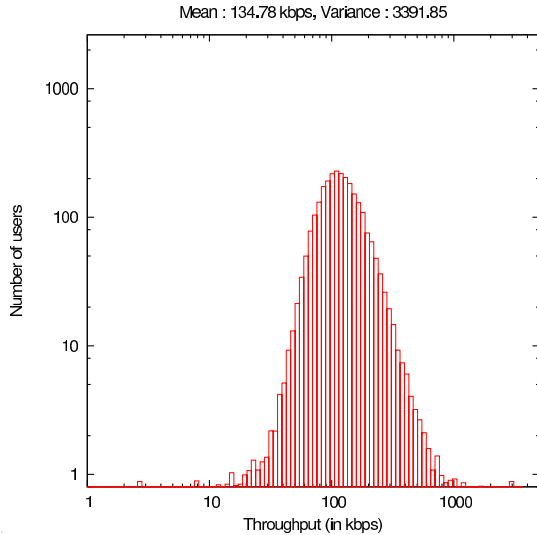
Fig. 7. Histogram of throughputs obtained by users for a mixed population where a half of users use closest association, and the other half Algorithm 4.

in an altruistic fashion as the one we investigate. We call such users selfish. Indeed they can be thought of as "intelligent users" which collect all the information needed by Algorithm 4 but choose to implement the following selfish variant:

Algorithm 5 *(Selfish User Deterministic Transition)*

Choose AP

$$a = \mathrm{argmin}_{a \in \mathcal{A}} \sum_{v \in \mathcal{U}_a} \frac{1}{f(\mathrm{SNR}(v))} \; .$$

Recall that the energy $\mathcal{E}_u(a)$ of a user associated with AP a, defined by Equation (6), comprises two terms. The first term is the potential delay (i.e. the inverse of the throughput) that user u would obtain in this cell. The second term represents the increase of potential delay experienced by other users of this cell if this user were to join. It could be the case that associating with a given AP a minimizes the sum of these two terms, whereas associating with a different AP b minimizes the first term alone. We call a user "selfish" if it is inclined to choose b.

To observe the performance of the system in the presence of selfish users, we perform three new experiments using the settings in §V-C.2. In the first experiment we assume all users are selfish. In the second experiment we assume one user is selfish. In the third experiment we assume a percentage (10%,50%,90%) of the users are selfish.

In the case of all-selfish wireless users we observe that the throughput distribution, shown in Figure 8(b), tends to be much narrower than the distribution of throughput achieved with association using Algorithm 4. This can be explained by the fact that a max-min fair rate is implemented in each cell, and that any lightly loaded cell will be selected by a selfish user, even if it does increase its performance by a very small amount and at the cost of a large decrease in the performance of others. This is confirmed by the association map that is shown in Figure 8(a). One can see that in a selfish scheme, users tend to associate to APs that may be far away to benefit from a lighter load. A direct consequence of the "selfish" scheme is that only a very small number of users see an improvement employing such algorithms (5% in our simulations), compared to the throughput they would have obtained with Algorithm 4. In addition, selfish users lead to a large number of users actually experiencing a 50% reduction in the throughput they would receive under a more fair scheme, like the one we propose. Essentially, if all users act selfishly they end up with less throughput than if they cooperate (Figure 8(c)).

This effect is also demonstrated in Figure 8(c), where we compare the throughput achieved by users when they all act selfishly and when they all adopt the proposed scheme. We notice that most data points fall below the diagonal indicating that there is a significant number of users that benefit from employing our scheme, while the number of users that experience smaller throughput is rather limited (the data points above the diagonal).

Note that the algorithm proposed in [10], where users associate to the AP that gives them the least expected time to transmit a data unit, may be thought as a "selfish" algorithm. In fact Algorithm 5 is a simplified version of this association decision, where the evaluation of the rate is made only on the downlink. Our simulations show that if a selfish mechanism is chosen as the default association option for 802.11 networks, the overall performance seen by the population of users might be degraded.

Nonetheless, an interesting question within this context concerns the potential benefit one client could gain if it were to act selfishly in an otherwise compliant set of APs and clients. Figure 9 presents this case. The y-axis gives the throughput that a user would obtain if it were the only new selfish client in
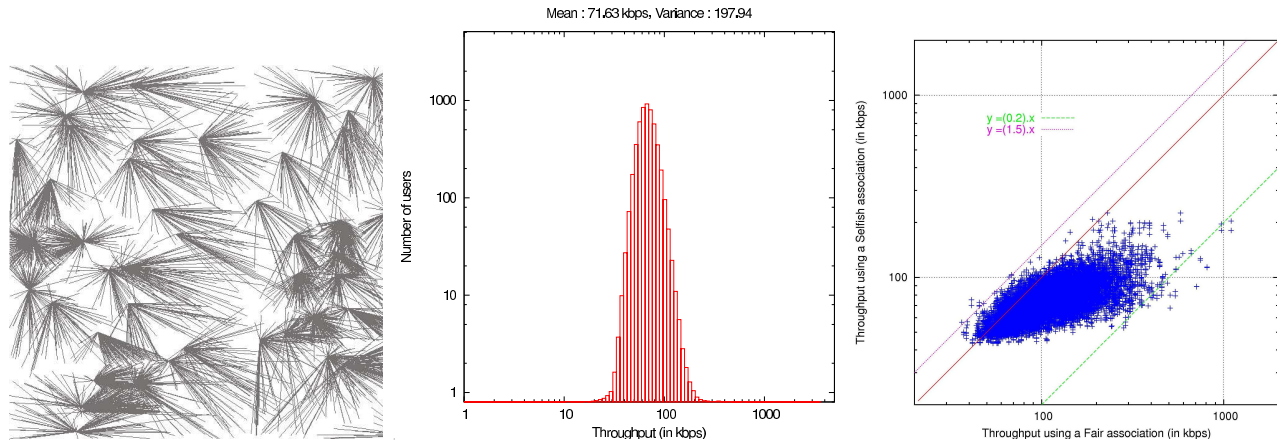
Fig. 8. Selfish association algorithms for sporadic topology: (a) map association (left), (b) histogram of the distribution of throughput (middle), (c) comparison between selfish and fair user throughput (right).

the system as a function of the throughput (x-axis) obtained when using Algorithm 4. The data set spanning the largest region, represented in light color, is the one where all other users are not misbehaving. We notice that indeed when one user has the choice of acting selfishly, it can achieve higher throughput than under the fair scheme. However, the gain is rather small: it reaches a throughput that is twice its fair value only in a few extreme cases. In many cases throughput achieved under a selfish scenario is even almost equivalent to the fair rate (48% of the cases). We note that the scenario we discuss essentially corresponds to the best scenario for a selfish user, since he/she is the only misbehaving user in the network. As one may expect, in the case of a mixed population, where other clients might have chosen to be selfish already, the individual gain from using a selfish algorithm is even smaller. The set of points represented with darker colors in Figure 9 shows the same comparison for a population which contains 10% and 50% of selfish users. For a population containing 50% of selfish users, shown in black, the gain is already marginal, only 26% of the users receive greater than their fair throughput.

VII. EXTENSIONS

In this section we relax some of the assumptions made during the presentation of our generic framework to demonstrate the flexibility of our solution. More specifically we first look into the potential for our algorithms to optimize the network state in terms of time fairness; we then investigate the use of overlapping frequencies, and lastly we consider the case of non-persistent traffic.
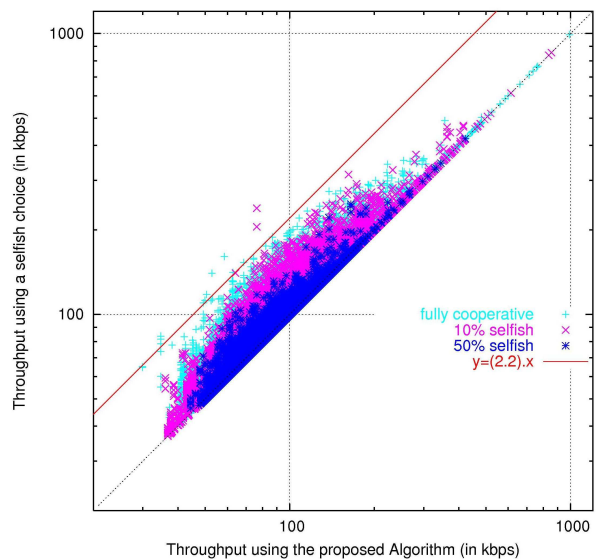

Fig. 9. Comparison of throughput obtained by a user deciding to adopt a selfish association algorithm.

A. Time Fairness vs. Rate Fairness

Several researchers have argued that the rate-based implementation of max-min fair bandwidth sharing in a 802.11 cell is not adapted to the wireless medium. Some have called this an anomaly. As pointed out for the first time by Heusse et al. in [6], and further explored by Radunovic et al. in [16], this mechanism is responsible for a sharp decrease in the performance experienced by all the users of a cell, when a single user joins under poor channel conditions. Such an observation led to a significant amount of research to recommend a max-min fair bandwidth sharing mechanism based not on rate but

on channel access time: the medium access time allocated to a client cannot be increased without decreasing that of a customer that was allocated less medium access time. Some mechanisms were proposed to achieve this objective in a single cell at the link layer (see e.g. [7]).

It turns out that our algorithm can be extended to this case. More precisely it can be extended to incorporate time fairness without modification.

Within this framework, user u, which associates with AP a with a time max-min fair share MAC, achieves a rate equal to

$$r_u = \frac{f(\text{SNR}(u))}{U_a} = \frac{f(\text{SNR}(u))}{\sum_{v \in \mathcal{U}} s(u,v)} \ . \tag{9}$$

so that we can rewrite the energy $\mathcal{E}$ as defined by (4):

$$\mathcal{E} = \sum_{u \in \mathcal{U}} \frac{1}{r_u} = \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{U}} \frac{s(u,v)}{f(\text{SNR}(u))} \ .$$

Our first observation is that even if the rate that each user obtains in a cell is different with a time-based bandwidth sharing, the global energy that defines the minimal potential delay allocation is unchanged. In particular it is straightforward to check that it is the same for a network where different cells may implement one or the other paradigm.

Similarly, the local energy function for a client u associated with a remains equal to:

$$\mathcal{E}(u) = \sum_{v \in \mathcal{U}; a(v)=a} \frac{1}{f(\text{SNR}(v))} + U_a \frac{1}{f(\text{SNR}(u))} \ .$$

The following remarkable observations can be made about it:

- If the cell associated with u and a implements a max-min fair rate share, the left term represents the potential delay of each client of this cell, whereas the right term represents the social cost of this client on this AP. The social cost is the sum of the increases of potential delay for other clients in this cell.
- If the cell associated with u and a implements a max-min fair time share, the opposite holds: the left term represents the social cost whereas the selfish cost of a client, given by its potential delay in this cell, is given by the right term.

This inversion of the social and selfish components of the same energy function could also create an incentive to play the social game of minimal potential delay fairness, as it does not require knowledge of the implementation, being based on rate or time. A selfish association scheme, by opposition, requires to know whether bandwidth is shared according to time or rate.

B. Overlapping Channels

In this paper we have assumed that channels are non interfering, so that two APs either fully interfere, if they choose the same channel, or dot not interfere at all. There are several good reasons to study a more general case: first, it was observed that in practice the performance degradation experienced on overlapping channels may not be detrimental [18]. Secondly, adding the possibility to use overlapping channels in a channel selection algorithm may provide a performance gain, as it increases channel diversity.

This case is actually easy to introduce in our algorithm. A first remark is that this has no impact on the user association algorithm, which remains unchanged. This is because the SNR measured directly by each user already contains interference coming from all channels, and it is therefore already included in the optimal selection that we have presented. For AP channel selection this case requires the introduction of a more general definition for $s(a,b)$: instead of taking value in $\{0,1\}$, this function takes value in the interval $[0;1]$, which translates the fraction of the power transmitted by b that is received by a when these two APs use overlapping channels. The values taken by this normalizing factor for 802.11b overlapping channels are presented in [13], where a distributed scheme is proposed for efficient channel assignment in wireless LANs, presented as a vertex coloring problem for weighted graphs.

C. Non-Persistent Traffic

So far, we have considered a scenario where users always have data to send. In this section, we first give the conditions under which the wireless link is the bottleneck and then analyze the case of a wireless access networks with non persistent users. Finally, we show that this case may be handled by a Gibbsian algorithm as well.

We assume in this section that the autorate function is continuous and linear. Hence the rate obtained

by a user is its associated SNR, multiplied by a normalizing constant. To avoid heavy notation in the following formula, we do not include this constant explicitly. In other words, the size of the unit of information used for the rate was chosen so that this constant is 1.

For $a \in \mathcal{A}$ and t a time slot, let $\text{act}(a, t)$ be equal to 1 if the cell associated with a is active during slot t and to 0 otherwise. For any $a \in \mathcal{A}$, we define the *long term activity factor* as:

$$\rho_a = \sum_{t=1}^{+\infty} \text{act}(a, t) .$$

During slot t, the SNR of user u associated with AP a is given by the following variant of (2):

$$\text{SNR}(u, t) = \frac{P_a(u)}{N_u + \sum_{b \in \mathcal{A} \,|\, b \neq a} s(b, a) P_b(u) \text{act}(b, t)} ,$$
$$(10)$$

Hence the time to transmit a unit of information to user u is given by:

$$\frac{1}{\text{SNR}(u, t)} = \frac{N_u}{P_a(u)} + \sum_{b \in \mathcal{A} \,|\, b \neq a} s(b, a) \text{act}(b, t) \frac{P_b(u)}{P_a(u)} .$$
$$(11)$$

We assume that data packets to be transmitted to u arrive to a according to some stationary point process with a finite intensity denoted by λ_u. We assume that the FIFO queue in the AP is stable (otherwise, we would be in the saturated case of the earlier sections), so that the transmission times of the packets of user u in AP a form a stationary point process with the same intensity. The workload introduced by user u in a is then equal to:

$$\lambda_u \mathbb{E}\left[\frac{1}{\text{SNR}(u, t)}\right] = \lambda_u \frac{N_u}{P_a(u)} \qquad (12)$$
$$+ \sum_{b \in \mathcal{A} \,|\, b \neq a} s(b, a) \rho_b \lambda_u \frac{P_b(u)}{P_a(u)} .$$

Summing on all users u associated with a, we obtain the following linear equation on the vector $\rho = (\rho_a)_{a \in \mathcal{A}}$:

$$\rho = \mathbb{A}\rho + B, \qquad (13)$$

where

$$\mathbb{A}_{a,b} = \begin{cases} \sum_{u \in \mathcal{U}} s(a, u) s(b, a) \lambda_u \frac{P_b(u)}{P_a(u)} & \text{for } b \neq a, \\ 0 & \text{for } b = a, \end{cases}$$
$$B_a = \sum_{u \in \mathcal{U}} s(a, u) \lambda_u \frac{N_u}{P_a(u)} .$$

The terms of this linear equation may be interpreted as follows: B_a is the sum of the durations needed to serve all users associated with a in the absence of interference; it is equal to ρ_a if either a is the unique AP or the only one operating on this channel. The coefficient $\mathbb{A}_{a,b}$ of the matrix is describing how much time needs to be spent in addition, when b is active, for AP a to serve all its users.

The solution of (13) is given by:

$$\rho = \mathbb{A}^* B, \quad \text{where } \mathbb{A}^* B = B + \mathbb{A}B + \mathbb{A}^2 B + \dots \ .$$
$$(14)$$

Note that $\mathbb{A}^*$ may have some infinite coefficients if the spectral radius of $\mathbb{A}$, that we denote $\text{spec}(\mathbb{A})$ is not smaller than 1. In this case, a finite solution for (13) cannot exist, proving that the network is always unstable.

If the spectral radius of $\mathbb{A}$ is smaller than 1, one can define formally a load vector with finite coefficient. Hence a wireless access network is stable/not saturated if and only if:

$$\begin{cases} \text{spec}(\mathbb{A}) < 1 \\ |\mathbb{A}^* B|_\infty < 1. \end{cases} \qquad (15)$$

Note that this condition depends on the association scheme used, as this is contained in the coefficients of $\mathbb{A}$, and B. Finding a good association mechanism that satisfies (15) is not easy in general. Intuitively the most stable user association is the one minimizing the norm $|\mathbb{A}^* B|_\infty$; solving this optimization problem implies analyzing the power series of matrix $\mathbb{A}$, and the maximum of certain sums of coefficients.

We introduce the following lower bound and upper bound vectors for ρ:

$$\check{\rho}_a \leq \rho_a \leq \tilde{\rho}_a, \qquad (16)$$

for all $a \in \mathcal{A}$, where

$$\check{\rho}_a = (B + \mathbb{A}B)_a$$
$$= \sum_{u \in \mathcal{U}_a} \lambda_u \frac{N_u}{P_a(u)}$$
$$+ \sum_{u \in \mathcal{U}_a, v \in \mathcal{U}_b \,|\, b \neq a} s(a, b) \lambda_u \lambda_v \frac{N_v}{P_b(v)} \frac{P_b(u)}{P_a(u)}$$
$$\tilde{\rho}_a = (B + \mathbb{A}\mathbf{1})_a$$
$$= \sum_{u \in \mathcal{U}_a} \lambda_u \frac{N_u}{P_a(u)} + \sum_{u \in \mathcal{U}_a \,|\, b \neq a} s(a, b) \lambda_u \frac{P_b(u)}{P_a(u)} .$$

It is interesting to observe that for all integers $p \leq 1$, the norm $|\check{\rho}|_p$, $|\tilde{\rho}|_p$ may be seen as an energy function which derives from a potential. Let us consider for example $|\check{\rho}|_1$; in the following formula we always denote by a (resp. b) the AP corresponding to user u (resp. v).

$$|\check{\rho}|_1 = \sum_{a \in \mathcal{A}} \check{\rho}_a = \sum_{u \in \mathcal{U}} \lambda_u \frac{N_u}{P_a(u)} +$$
$$\sum_{\{u,v\} \subseteq \mathcal{U}, a \neq b} s(a,b) \lambda_u \lambda_v \left(\frac{N_v}{P_b(v)} \frac{P_b(u)}{P_a(u)} + \frac{N_u}{P_a(u)} \frac{P_a(v)}{P_b(v)} \right) . \tag{17}$$

It is thus possible for users to choose their association in a distributed manner based on the minimization of $|\check{\rho}|_1$, as long as they are able to compute their own local energy given in this case by:

$$\mathcal{G}_u = \lambda_u \frac{N_u}{P_a(u)} \left(1 + \sum_{b \neq a} \sum_{v \in \mathcal{U}_b} s(a,b) \lambda_v \frac{P_a(v)}{P_b(v)} \right)$$
$$+ \lambda_u \sum_{b \neq a} s(a,b) \frac{P_b(u)}{P_a(u)} \sum_{v \in \mathcal{U}_b} \lambda_v \frac{N_v}{P_b(v)} . \tag{18}$$

The terms of this energy can be computed and communicated in an efficient manner: we assume that user u associated with AP a measures the ratios $\lambda_u N_u / P_a(u)$ and $\lambda_u P_b(u)/P_a(u)$ for all APs b operating on the same channel. These two variables are reported to a, which computes the cumulated sum of these variables taken on all users in its cell. AP a reports the value of these sums to any AP designated by one of its users. When a user needs to make a transition, it can compute its local energy based on these values that are maintained and advertised by each AP. The same type of algorithm can be thought of for $\tilde{\rho}$ and for any norm.

VIII. Conclusion

The fast and unmanaged deployment of wireless LANs makes the need for self-configuration capabilities on wireless devices of utmost importance. There is an imminent need for distributed algorithms that will allow both WLAN operators and users to extract as much benefit as possible from the shared spectrum, adjusting to network conditions and efficiently sharing the wireless resources.

We have proposed two self-configuration algorithms that can be deployed today in 802.11 Access Points and clients. These algorithms use local measurements in order to (1) allow APs to select a channel that will experience minimal interference with neighboring APs and (2) give WLAN clients a fair share of the global resource (i.e. the total capacity of the network created by all the interfering wireless LANs). We have shown that our algorithms are stable under various conditions and that they perform better than currently used techniques. Moreover, they are flexible enough to incorporate effects such as shadowing, the use of overlapping channels, and various traffic hypotheses. We discussed scenarios of incremental deployment and showed that clients and APs using our algorithms will not have a handicap in non cooperative environments, while the incentives to misbehave remain small. Implementation of the proposed algorithms relies on simple software modifications that could be incorporated within the efforts of the IEEE 802.11k task group. It allows us to make more generally several recommendations for the design of self organized wireless networks. In particular we demonstrate in this paper that implementing distributed decision which takes into account the individual gain as well as the social cost could easily lead to much higher efficiency of spectrum use, and improved performance for a large class of users.

In addition to our extensive simulations, we now need to implement the proposed algorithms and deploy them on real-life testbeds. More analysis is also needed to understand the performance of these algorithms when APs belong to different operators (such as hot-spot networks). We further intend to study extensions to support wireless meshed networks.

References

[1] Aditya Akella, Glenn Judd, Peter Steenkiste, and Srinivasan Seshan. Self management in chaotic wireless deployments. In *MobiCom '05: Proceedings of the 11th annual international conference on Mobile computing and networking*, 2005.

[2] Yigal Bejerano, Seung-Jae Han, and Li (Erran) Li. Fairness and load balancing in wireless LANs using association control. In *MobiCom '04: Proceedings of the 10th annual international conference on Mobile computing and networking*, pages 315–329, 2004.

[3] P. Brémaud. *Markov Chains, Gibbs Field, Monte Carlo Simulation and Queues*. Springer-Verlag, 1999.

[4] Sunwoong Choi, Kihong Park, and Chong kwon Kim. On the performance characteristics of wlans: revisited. In *SIGMETRICS '05: Proceedings of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, pages 97–108, 2005.

[5] Violeta Gambiroza, Bahareh Sadeghi, and Edward W. Knightly. End-to-end performance and fairness in multihop wireless backhaul networks. In *MobiCom '04: Proceedings of the 10th annual international conference on Mobile computing and networking*, pages 287–301, 2004.

[6] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda. Performance anomaly of 802.11b. In *Proceedings of IEEE INFOCOM'03 (The Conference on Computer Communications)*, 2003.

[7] M. Heusse, F. Rousseau, and A. Duda. Idle sense: An optimal access method for high throughput and fairness in rate diverse wireless LANs. In *Proceedings of the 2005 conference on Applications, technologies, architectures, and protocols for computer communications*, 2005.

[8] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu. Impact of interference on multi-hop wireless network performance. In *MobiCom'03: Proceedings of the 9th annual international conference on Mobile computing and networking*, pages 66–80, 2003.

[9] Murali Kodialam and Thyaga Nandagopal. Characterizing achievable rates in multi-hop wireless networks: the joint routing and scheduling problem. In *MobiCom '03: Proceedings of the 9th annual international conference on Mobile computing and networking*, pages 42–54, 2003.

[10] T. Korakis, O. Ercetin, S. Krishnamurthy, L. Tassiulas, and S. Tripathi. Link quality based association mechanism in IEEE 802.11h compliant wireless LANs. In *Proceedings of the first workshop on Resource Allocation in Wireless NETworks (RAWNET'05)*, April 2005. (available at: `http://www.rawnet.org/rawnet_korakis.pdf`).

[11] B.-J. Leung, K.K.; Kim. Frequency assignment for IEEE 802.11 wireless networks. In *Proceedings of the 58th IEEE Vehicular Technology Conference (VTC 2003-Fall)*, volume 3, pages 1422–1426, oct 2003.

[12] Laurent Massoulié and James Roberts. Bandwidth sharing: objectives and algorithms. *IEEE/ACM Trans. Netw.*, 10(3):320–328, 2002.

[13] Arunesh Mishra, Suman Banerjee, and William Arbaugh. Weighted coloring based channel assignment in WLANs. *under submission for Mobile Computing and Communications Review*, 2005.

[14] C. Na, J. Chen, and T. Rappaport. Measured traffic statistics and throughput of ieee 802.11b public wlan hotspots with three different applications. *submitted to IEEE Transactions on Wireless Communications*, 2005.

[15] Lili Qiu, Ranveer Chandra, Kamal Jain, and Mohammad Mahdian. On the placement of integration points in multi-hop wireless networks. In *Proceedings of IEEE International Conference on Network Protocols*, 2004.

[16] B. Radunovic and J. Y. Le Boudec. Rate performance objectives of multi-hop wireless networks. In *Proceedings of IEEE International Conference on Computer Communications*, March 2004.

[17] A. Raniwala and T. Chiueh. Architecture and algorithms for an IEEE 802.11-based multi-channel wireless mesh network. In *Proceedings of IEEE International Conference on Computer Communications*, March 2005.

[18] J. Robinson, K. Papagiannaki, C. Diot, X. Guo, and L. Krishnamurthy. Experimenting with a multi-radio mesh networking testbed. In *1st workshop on Wireless Network Measurements (WiNMee)*, April 2005.