Poster: Optimal Load-Balancing for High-Density Wireless Networks with Flow-Level Dynamics

Bin Li
Department of ECBE
University of Rhode Island
Kingston, Rhode Island, USA
binli@uri.edu

Xiangqi Kong
Department of ECBE
University of Rhode Island
Kingston, Rhode Island, USA
xqkong@uri.edu

Lei Wang
School of Software
Dalian University of Technology
Dalian, Liaoning, China
lei.wang@dlut.edu.cn

ABSTRACT

We consider the load-balancing design for forwarding incoming flows to access points (APs) in high-density wireless networks with both channel fading and flow-level dynamics, where each incoming flow has a certain amount of service demand and leaves the system once its service request is complete. The efficient load-balancing design is strongly needed for supporting high-quality wireless connections in high-density areas. In this work, we propose a Joint Load-Balancing and Scheduling (JLBS) Algorithm that always forwards the incoming flows to the AP with the smallest workload in the presence of flow-level dynamics and each AP always serves the flow with the best channel quality. Our analysis reveals that our proposed JLBS Algorithm not only achieves maximum system throughput, but also minimizes the total system workload in the heavy-traffic regime. Moreover, we observe from both our theoretical and simulation results that the mean total workload performance under the proposed JLBS Algorithm does not degrade as the number of APs increases, which is strongly desirable in high-density wireless networks.

CCS CONCEPTS

• **Networks → Network resources allocation**; **Wireless local area networks**;

KEYWORDS

Wireless networks, flow-level dynamics, load-balancing, throughput, mean delay, heavy-traffic analysis

ACM Reference Format:
Bin Li, Xiangqi Kong, and Lei Wang. 2018. Poster: Optimal Load-Balancing for High-Density Wireless Networks with Flow-Level Dynamics. In *Mobihoc '18: The Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing, June 26–29, 2018, Los Angeles, CA, USA.* Los Angeles, CA, USA, 2 pages. https://doi.org/10.1145/3209582.3225205

1 PROBLEM SETUP

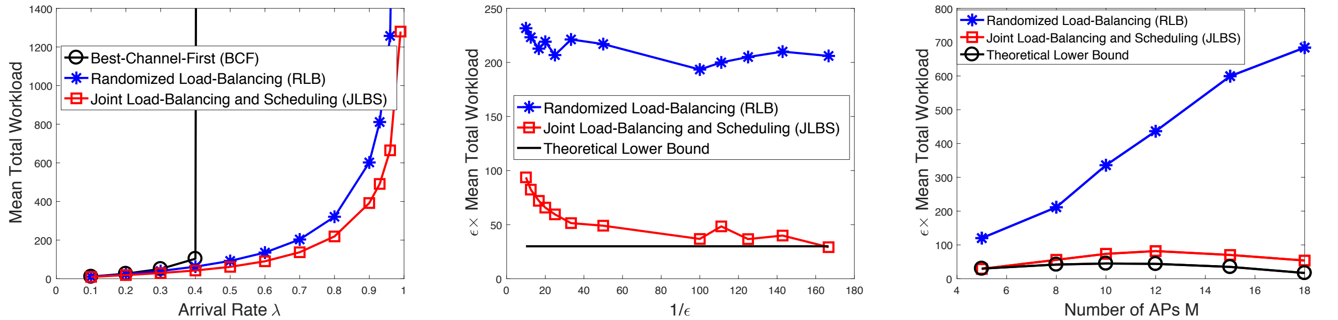
Multiple access points (APs) must be deployed for providing satisfactory services for users in high-density areas, such as convention centers, auditoriums, hotel meeting rooms, lecture halls, sports stadiums, and concert halls. However, in traditional wireless local access networks (WLANs), each user is automatically associated with the AP that has the best channel quality, which causes significant load imbalance among APs and results in poor network performance. This raises a natural question in how to develop an efficient joint load-balancing and scheduling algorithm that first determines which AP an incoming user should associate with, and then each individual AP needs to decide which users it serves. The goal of such an algorithm is to maximize system throughput (or equivalently support network users as many as possible) and to minimize average user's delay.

To this end, in this paper, we consider a wireless network with M APs operating in orthogonal channels. We assume that the system operates in a *slotted time* manner. Due to the wireless interference, within each AP, at most one flow can be served in each time slot. Let $A_\Sigma[t]$ denote the number of flows arriving at the system in time slot t that is bounded and independently and identically distributed (i.i.d.) over time with mean $\lambda_\Sigma > 0$. We use $F_j[t]$ to denote the number of packets of newly arriving flow j that follows any probability distribution with finite support. We use $N_m[t]$ to denote the number of flows in AP m in time slot t. We also use $\mathcal{A}_\Sigma[t]$ and $\mathcal{N}_m[t]$ to denote the set of newly arriving flows at the system and the set of existing flows in AP m in time slot t, respectively. Let $R_j[t]$ be the number of residual packets of flow j in time slot t. We assume that each AP has a finite number of possible channel rates with $c_{\max}$ denoting its maximum channel rate. We use $C_{m,j}[t]$ to capture wireless channel fading of each flow j in the m^{th} AP, which measures the maximum number of packets that can be transmitted in time slot t if flow j is scheduled. We assume that $(C_{m,j}[t])_{j\in\mathcal{N}_m[t]}$ are independently distributed across APs and i.i.d. over both time and flows within each AP.

In order to characterize the underlying dynamics of flows, we introduce following notations. Let $W_m[t] \triangleq \sum_{j\in\mathcal{N}_m[t]} \lceil R_j[t]/c_{\max} \rceil$ be the total workload in AP m in time slot t that measures the minimum number of slots required for completing all existing service requests in AP m. We use $\nu_\Sigma[t] \triangleq \sum_{j\in\mathcal{A}_\Sigma[t]} \lceil F_j[t]/c_{\max} \rceil$ and $\nu_m[t] \triangleq \sum_{j\in\mathcal{A}_m[t]} \lceil F_j[t]/c_{\max} \rceil$ to denote the total amount of new workload arriving at the system and the amount of new workload injected to AP m under some load-balancing policy in time slot t, respectively, where $\mathcal{A}_m[t]$ denotes the set of arriving flows at AP m in time slot t. We also use $A_m[t]$ to represent the number of newly arriving flows at AP m in time slot t. Let $\rho \triangleq \mathbb{E}[\nu_\Sigma[t]] = \lambda_\Sigma w$ be the traffic intensity, where $w \triangleq \mathbb{E}\left[\lceil F_j[t]/c_{\max} \rceil\right]$ denotes the expected minimum number of slots required for serving a newly arriving

Mobihoc '18, June 26–29, 2018, Los Angeles, CA, USA

ACM ISBN 978-1-4503-5770-8/18/06…$15.00
https://doi.org/10.1145/3209582.3225205



(a) Throughput performance **(b) Heavy-traffic performance** **(c) Impact of number of APs**

Figure 1: Performance of the JLBS Algorithm

flow. We define $\mu_m[t]$ to be the amount of workload decreasing at AP m in time slot t.

2 LOAD-BALANCING ALGORITHM DESIGN

Based on the setup described in Section 1, the evolution of the workload $W_m[t]$ at each AP m can be characterized as follows: $W_m[t+1] = W_m[t] + v_m[t] - \mu_m[t], \forall m = 1, \ldots, M$. We call AP m *stable* if its average workload is finite. We say that the system is stable if all its APs are stable. The *capacity region* Λ is defined as a maximum set of traffic intensity ρ for which the system is stable under some policy. It is shown in our technical report [1] that $\Lambda = \{\rho : \rho \leq M\}$, where we recall that M is the number of APs. Next, we propose a joint load-balancing and scheduling algorithm.

Joint Load-Balancing and Scheduling (JLBS) Algorithm: In each time slot t, given the workload $\mathbf{W}[t] = (W_m[t])_{m=1}^M$, perform (1) *Load-balancing decision*: Forward all the arriving flows to the AP with the smallest workload, i.e.,

$$\mathbf{A}^*[t] \in \operatorname*{argmin}_{\mathbf{A}=(A_m)_{m=1}^M \geq \mathbf{0}: \sum_{m=1}^M A_m = A_\Sigma[t]} \langle \mathbf{A}, \mathbf{W}[t] \rangle. \tag{1}$$

(2) *Scheduling decision*: Within each AP m, serve the flow j_m^* with the maximum channel rate among all its existing flows, i.e.,

$$j_m^* \in \operatorname*{argmax}_{j \in \mathcal{N}_m[t]} C_{m,j}[t]. \tag{2}$$

Next, we show that our proposed JLBS Algorithm achieves both throughput-optimality and heavy-traffic optimality. The detailed proofs can be found in our technical report [1].

Proposition 1. *The JLBS Algorithm is throughput-optimal, i.e., it stabilizes the system for any traffic intensity lying strictly inside the capacity region* Λ.

To characterize the heavy-traffic performance of the JLBS Algorithm, we consider the workload arrival process $\{v_\Sigma^{(\epsilon)}[t]\}_{t\geq 0}$, parameterized by $\epsilon > 0$, with traffic intensity $\rho^{(\epsilon)}$ satisfying $\epsilon = M - \rho^{(\epsilon)} > 0$ and $\text{Var}(v_\Sigma^{(\epsilon)})$. Here, ϵ characterizes the closeness of the traffic intensity to the boundary of the capacity region, and is usually referred as *heavy-traffic parameter.*

Proposition 2. *The JLBS Algorithm is heavy-traffic optimal in the sense that it minimizes the total system workload in the heavy-traffic limit, i.e.,* $\epsilon \downarrow 0$.

3 SIMULATION RESULTS

In this section, we perform extensive simulations to validate the efficiency of our proposed JLBS Algorithm by comparing it with the Best-Channel-First (BCF) Algorithm and the Randomized Load-Balancing (RLB) Algorithm. Here, both BCF and RLB Algorithms use the same scheduling decision as the JLBS Algorithm under which each AP always schedules the flow with the best channel quality. However, for the load-balancing decision, the BCF Algorithm always forwards the arriving flows to the AP with the best signal quality, while the RLB Algorithm makes the load-balancing decision in a purely randomized fashion. We assume that the number of flows arriving at the system in each time slot follows a Bernoulli distribution with mean λ. Each flow at each AP faces i.i.d. channel fading with rates $0, 1, 5, 10$ and corresponding probability $0.1, 0.2, 0.5, 0.2$. The flow size F is equal to $10 \times \beta$ with probability $(w-1)/(\beta-1)$ and 10 otherwise. We let w be equal to the number of APs M and thus the capacity region Λ is $\{\lambda : 0 < \lambda \leq 1\}$. We set $M = 5$ and $\beta = 20$ in the simulations.

We can see from Fig. 1b that the mean total workload under the JLBS Algorithm converges to the theoretical lower bound derived in our technical report [1, Proposition 3], while the RLB Algorithm always keeps it away from the theoretical lower bound. This confirms the heavy-traffic optimality of the JLBS Algorithm, i.e., it minimizes the mean total workload as the heavy-traffic parameter ϵ diminishes. Fig. 1c illustrates that the performance of the JLBS Algorithm stays close to the theoretical lower bound and its performance does not scale with the number of APs. This desired property indicates that our proposed JLBS Algorithm works perfectly in the high-density wireless networks.

ACKNOWLEDGEMENTS

This work is supported in part by US NSF grant CNS-1717108, China NSF grant 61733002 and Tianjin Key Laboratory of Advanced Networking (TANK).

REFERENCES

[1] B. Li, X. Kong, and L. Wang. 2018. Optimal Load-Balancing for High-Density Wireless Networks with Flow-Level Dynamics. (2018). Technical Report. Available at https://arxiv.org/pdf/1805.00158.pdf.