This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2919590, IEEE Transactions on Communications

Joint Deployment and Pricing of Next-Generation WiFi Networks

Konstantinos Poularakis[‡], George Iosifidis[*], and Leandros Tassiulas[‡]

[‡]Department of Electrical Engineering, and Institute for Network Science, Yale University, USA
[*]School of Computer Science and Statistics, Trinity College Dublin, Ireland

***Abstract*—WiFi is increasingly used by carriers for opportunistically offloading the cellular network infrastructure or even for increasing their revenue through WiFi-only plans and WiFi on-demand passes. Despite the importance and momentum of this technology, the current deployment of WiFi access points (APs) by the carriers follows mostly a heuristic approach. In addition, the prevalent free-of-charge WiFi access policy may result in significant opportunity costs for the carriers as this traffic could yield non-negligible revenue. In this paper, we study the problem of optimizing the deployment of WiFi APs and pricing the WiFi data usage with the goal of maximizing carrier profit. Addressing this problem is a prerequisite for the efficient integration of WiFi to next-generation carrier networks. Our framework considers various demand models that predict how traffic will change in response to alteration in price and AP locations. We present both optimal and approximate solutions and reveal how key parameters shape the carrier profit. Evaluations on a dataset of WiFi access patterns indicate that WiFi can indeed help carriers reduce their costs while charging users about 50% lower than the cellular service.**

***Index Terms*—WiFi deployment, WiFi pricing, Profit optimization, Network design**

I. Introduction

A. Motivation

Mobile data traffic continues to grow, rising by more than 79% worldwide in 2018 [2]. This "data tsunami" threatens to drain the capacity of cellular networks resulting in degradation of user service quality [3]. Traditional network capacity expansion methods, such as cellular technology upgrades and additional spectrum acquisition, are expensive and constantly outpaced by the continuing traffic increase. Therefore, carriers are today in a position where their traffic and hence expenditures increase sharply, but revenues do not necessarily follow[1].

This demand has motivated many new business models that leverage the availability of WiFi to complement or increase cellular network capacity [5]. For example, Google recently announced a hybrid mobile data plan, called Project Fi [6], where user devices can automatically access either of three partner cellular networks (T-Mobile, Sprint and US Cellular) or any available open WiFi network. Users are not charged for WiFi, while they pay $\$10$ per gigabyte of cellular data usage. Additional plans that bundle WiFi and cellular data are currently advertised by other small virtual carriers. For example, Republic Wireless charges WiFi data alone at the flat rate of $5, while the price increases up to $\$45$ when bundled with caps of cellular data (provided by Sprint's cellular network) and other services [7], [8]. Similarly, US Mobile and Boingo offer cheap prepaid WiFi-only plans (at $\$10$ flat rate), especially suitable for users who travel frequently or live a domestic lifestyle [9], [10]. Comcast advertises WiFi on-demand passes to millions of access points (APs) that are free for its subscribers but are priced from $\$3.95$ per two hours to $\$54.95$ per month for non-subscribers [11].

The existence of so many models and solutions is not surprising since there are several reasons favoring the use of WiFi by carriers today. First, WiFi utilizes unlicensed spectrum which is separate from licensed cellular bands. In fact, Federal Communications Commission recently released an additional 100 megahertz of spectrum for WiFi usage [12]. Second, the latest WiFi technology can provide a very high data rate. Third, WiFi and cellular technology are continuously converging with the development of mechanisms and protocols following the specifications of the latest 3GPP releases [13], while this integration is expected to be even higher in next-generation networks [14].

To reap the potential benefits from WiFi technology, it is required to incorporate it properly in cellular networks. Indeed, many carriers today deploy their own WiFi APs to offload traffic from cellular networks. To ease deployment and keep costs low, carriers often cooperate with the owners of crowded places such as shopping malls, restaurants and cafes on AP deployment[2]. When such venues are not available or cooperation is not possible, carriers have the option to build 'from the ground up' new APs[3]. In this case, however, the AP deployment cost increases to account for labor expenses, cost of backhaul installation, power supply, etc.

The current deployment of WiFi APs by the carriers is accomplished mainly in an ad hoc fashion using simple heuris-

Part of this work appeared in the proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc), 2016 [1]. This work was supported partly by the Army Research Office under Agreement Number W911NF-18-10-378, the National Science Foundation under Grant CNS 1815676 and the Science Foundation Ireland under Grant 17/CDA/4760. K. Poularakis acknowledges the Bodossaki Foundation, Greece, for a postdoctoral fellowship.

[1]As an example, China Mobile reported for 2014 115.1% increase in mobile data traffic and, at the same time, 10.2% profit reduction [4].

[2]For example, AT&T has been cooperating with hotspot venues (such as Starbucks) to install its own public WiFi APs [15].

[3]For example, see Facebook's Express WiFi initiative targeting many underserved areas around the world [16].

tic approaches (e.g., deployment to the densest areas and/or venues where installation costs are the lowest). Therefore, from a network perspective the resulting architecture is not optimized. Additionally, it is not clear how much the mobile users should be charged for using these auxiliary networks. Charging cheaper the WiFi than the cellular access can be regarded by the users as a discounted price and an incentive to offload their data. However, the currently prevalent free-of-charge WiFi access policy may result in significant opportunity costs for the carriers as this traffic could yield non-negligible revenue because of the users who would otherwise have consumed their cellular data plans (opportunity cost).

It is clear from the above that there are currently many key questions which we need to ask in a systematic way and answer in a rigorous fashion, so as to be able to unleash the power of the next-generation carrier-grade WiFi data offloading solutions. In particular:

1) Which AP deployment policy maximizes the carrier profit, i.e., the difference between revenue and costs?
2) How does the pricing policy affect the user demand for WiFi and cellular data, and how can it be determined to reduce opportunity costs but not offloading benefits?

The above questions remain open, since until now WiFi data offloading has been an opportunistic solution. To answer these questions, we need a valid mechanism to identify a set of AP locations and prices, which takes into account the deployment costs and the preferences of the users. However, users are often unpredictable which makes difficult to model their demand for WiFi and cellular data. Therefore, we need to investigate different models for user demand, to ensure that out analysis covers a wide spectrum of cases.

B. Methodology and contributions

In this paper we follow a systematic methodology in order to overcome these obstacles and answer the above questions. Motivated by new business models that leverage WiFi to complement or increase cellular network capacity, we focus on a virtual carrier like Google, Republic Wireless and US Mobile. This carrier has the option to deploy WiFi APs to reduce usage of other providers' cellular networks and save the respective payments. For the users, we describe three alternative models to capture their demand at the WiFi APs. Namely, we consider the *Maximum-Surplus Demand (MSD)*, *Constant Elasticity Demand (CED)* and *Logit Demand (LD)* models which have been also used in studies of Internet transit [17] and residential broadband [18] markets. For each demand model, we formulate the *Joint WiFi AP Deployment and Pricing (JWDP)* problem with the goal of maximizing the carrier's profit.

The complexity of the JWDP problem highly depends on the demand model. For the MSD model, we show that the optimal WiFi access price belongs to a finite set of values, and derive an approximation algorithm using a connection to a *facility location-type problem* [19]. We then present an optimal solution in closed-form for the CED model, and show that *the optimal number of deployed APs is a decreasing function of the AP deployment cost, the net carrier profit for delivering data by the cellular infrastructure, the cost for delivering data via WiFi APs, and the "sensitivity" of the users to price changes*. Finally, for the LD model we present a randomized greedy based approximation algorithm by expressing the problem as the maximization of a *submodular set function* [20]. These solutions reveal the dependency of carrier profit on key system parameters and provide insights to carriers or policy makers.

We evaluate the performance of the proposed solutions using WiFi access patterns collected from real users. Referencing cost and pricing information from several sources, we synthesize a variety of scenarios and investigate the impact of WiFi in each scenario. The results we present clearly differentiate situations in which WiFi can notably help carriers, and those in which the gains are marginal. The main contributions of this work can be summarized as follows:

- *Joint WiFi Deployment and Pricing*. We introduce the JWDP problem that jointly derives WiFi AP deployment and pricing policies to maximize the carrier's profit, using general and diverse models of user demand.
- *Analytical and approximate solutions*. We present an optimal solution to the JWDP problem for the CED model. For the MSD and LD models, we present approximate solutions by identifying non-trivial connections to facility location and submodular function optimization problems respectively.
- *Dataset-driven evaluation*. We use a dataset of WiFi access patterns collected from real users and examine the cost-benefit impact of our solutions in different scenarios. We find that WiFi can help carriers reduce their costs, while charging users about 50% lower than the cellular service. Some small virtual carriers may even halve their payments for using other's cellular infrastructure. However, deploying APs in an ad-hoc manner can lead to a net loss for the carrier.

The rest of the paper is organized as follows: Section II describes the system model and defines the JWDP problem formally. In Section III, we present the optimal and approximate solutions for the different demand models. We study practical issues including the congestion of the APs in Section IV. Section V presents our dataset-driven evaluation results, while Section VI reviews our contribution compared to related work. We conclude our work in Section VII.

II. Model and problem formulation

A. Carrier model

We consider a virtual carrier (e.g., Google, Republic Wireless, US Mobile) that provides mobile data services in a certain geographical area, as in Figure 1. The virtual carrier has to pay its cellular provider (e.g., Sprint) for the data sent over its network. The real payment information is proprietary, but typically the billing is per unit of data [21]. We denote this cost by c_b (\$/GB). The carrier charges users for the service at the rate of $p_b > 0$ (\$/GB). For example, Google charges $\$10$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2919590, IEEE Transactions on Communications

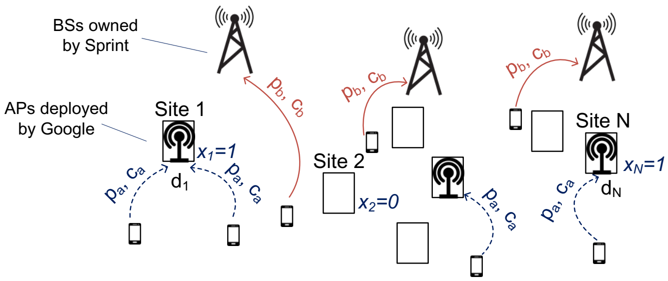


Fig. 1. Graphical illustration of the proposed model. Rectangles represent candidate sites for WiFi AP deployment. Arrows indicate communication links between users and WiFi APs or cellular base stations (BSs).

for each GB of data consumed to the nearest 0.01GB [6].[4] To ensure the viability of the carrier, it should be $p_b > c_b$.

The carrier may deploy several WiFi APs to divert traffic from the cellular network and reduce its costs. In practice, the deployment will be constrained to a finite collection $\mathcal{N}$ of N candidate sites. For example, APs can be deployed in venues their owners have agreed to cooperate with the carrier on AP installation or in sites the carrier has acquired or leased for building new APs. The deployment cost will be different depending on the scenario since different cost factors are involved, e.g. wireless equipment purchase and installation, backhaul link connection, site rental, labor cost, repair/maintenance cost, etc. We note that some of these costs represent capital investments (CAPEX) paid once while others are operational costs (OPEX) paid multiple times, e.g., on a monthly basis. To capture both types of cost, we take the aggregate of the OPEX cost and the CAPEX cost amortized for a period of study (one month). Then, we denote by $d_n > 0$ the monthly cost associated with the deployment of an AP at site $n \in \mathcal{N}$.

The carrier requests a payment $p_a > 0$ (\$/GB) from the users for the WiFi access, where p_a is not necessarily equal to p_b. Typically, p_a will be lower than p_b (or even zero), which can be regarded as a discounted price or an incentive for the users to shift their data from the cellular to the WiFi network. For example, by setting p_a equal to $p_b/2$, the carrier can advertise this as billing only half of the data consumption if the user voluntarily subscribes to a hybrid cellular/WiFi data plan. We refer by p_a the *pricing policy* of the carrier. Transmitting data by an AP causes a servicing cost (e.g., due to energy consumption), denoted by $c_a > 0$ (\$/GB), which is typically lower than the cellular servicing cost ($c_a < c_b$).

We introduce the variable $x_n \in \{0, 1\}$ to indicate whether an AP at site n is deployed ($x_n = 1$) or not ($x_n = 0$). These variables constitute the *deployment policy* of the carrier:

$$\boldsymbol{x} = (x_n \in \{0, 1\} : n \in \mathcal{N}) \; . \tag{1}$$

We also denote by $\mathcal{N}_{\boldsymbol{x}} \subseteq \mathcal{N}$ the subset of sites where APs are deployed, i.e., containing every site n with $x_n = 1$.

[4]Our work is focused on the volume-based pricing scheme where every unit of data consumed is charged. Alternative schemes included time-dependent pricing, data-cap pricing, bundled pricing of multiple services, etc.

B. User demand models

In response to the *deployment policy ($\boldsymbol{x}$)* and the *pricing policy (p_a)*, a portion of the user traffic will be directed to the APs instead of the cellular base stations (BSs). We denote by $Q_n(\boldsymbol{x}, p_a)$ the respective demand for the AP at site n. Similarly, $Q_b(\boldsymbol{x}, p_a)$ specifies the demand for cellular BSs. Intuitively, the lower the AP price p_a and the higher the number of deployed APs are, the more demand is offloaded to the WiFi network (increasing $Q_n(\boldsymbol{x}, p_a)$, $\forall n$) and the lower the cellular network load becomes (decreasing $Q_b(\boldsymbol{x}, p_a)$). Nevertheless, uncovering the exact shape of the demand functions is challenging. To ensure that our analysis covers a wide spectrum of users, we consider three alternative demand models: namely (i) *maximum-surplus demand*, (ii) *constant elasticity demand* and (iii) *logit demand* models. Each demand model comes with its pros and cons as we explain below.

i. Maximum surplus demand (MSD). The MSD model has been widely used in the economics and marketing literature [22]. In this model, a seller offers a variety of products and each consumer chooses a single product that yields the greatest nonnegative *surplus*. The consumer surplus for a product is defined as the difference between the consumer reservation price (or the maximum monetary value she is willing to pay for this product) and the price set by the seller. In the context of mobile data service, the consumers are the subscribers of the carrier (or users) and the choice is between the WiFi and cellular networks.

Formally, we consider a set $\mathcal{K}$ of K mobile users. The users are arbitrarily distributed within the area based on a probability distribution that can be estimated by the carrier (e.g., by analyzing historical location data). A user $k \in \mathcal{K}$ can associate only to APs in communication range, where $\mathcal{N}_k \subseteq \mathcal{N}$ denotes the respective set of sites. Without loss of generality, we consider all users to have the same demand for data[5], normalized to 1 (GB). A user k is willing to pay up to $l_{ka} \geq 0$ ($\$/GB$) for downloading data by a WiFi AP and up to $l_{kb} \geq 0$ ($\$/GB$) by a cellular BS. The surplus of user k for an AP is $l_{ka} - p_a$, while it is $l_{kb} - p_b$ for a BS. We consider that $l_{kb} - p_b \geq 0$, otherwise the user would have switched to a different carrier.

User k will be served by a WiFi AP instead of a cellular BS if this yields higher surplus, i.e., if there exists an AP $n \in \mathcal{N}_k$ such that $x_n = 1$ and $l_{ka} - p_a \geq l_{kb} - p_b$. Otherwise, user k will be served by a BS. When more than one APs cover the user, we pick the one that is closer to the user. We will revisit the closest AP selection assumption in Section IV where the impact of AP congestion will be studied. We denote with $n(k) \in \mathcal{N}_k \cup \{b\}$ the choice of user k based on the above rule, where b indicates the cellular network. Then, the demand of AP n can be written as:

$$Q_n(\boldsymbol{x}, p_a) = \sum_{k \in \mathcal{K}} 1_{\{n(k)=n\}}, \tag{2}$$

[5]This is a mild assumption as users generating more demand than others can be represented by multiple users in the $\mathcal{K}$ set.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2919590, IEEE Transactions on Communications

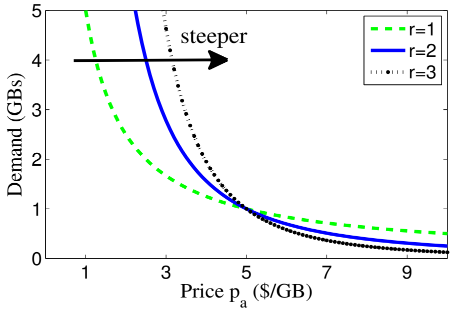


Fig. 2. CED function example. The demand of an AP decreases with price, with the function becoming steeper with r.

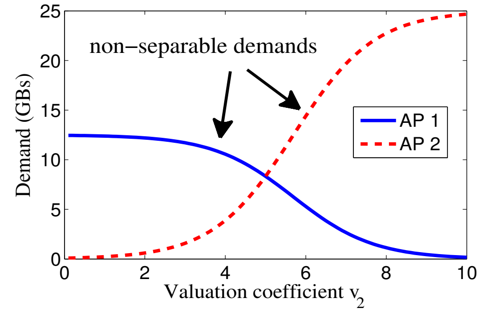


Fig. 3. LD function example. The two APs "compete" for serving the same demand.

where $1_{\{.\}}$ is the indicator function, i.e., it is equal to one if the condition in the subscript is true; otherwise zero. Similarly, we can write the demand of the cellular network as:

$$Q_b(\boldsymbol{x}, p_a) = \sum_{k \in \mathcal{K}} 1_{\{n(k)=b\}}, \tag{3}$$

Although plausible, the MSD model has a limitation; it assumes that the location ($\mathcal{N}_k$ sets) and the amounts of money that each user is willing to pay for WiFi and cellular data (l_{ka} and l_{kb} values) are known in advance. The latter requires the estimation of $2K$ values, 2 values for each user. These values essentially represent a distribution of demand in the network the carrier needs to estimate. Accurate estimation may be possible in regions with small population for which the carrier can perform extensive surveys to gather the required information. For large population of users, however, alternative more scalable demand models are needed, as we discuss below.

ii. Constant elasticity demand (CED). The CED model is derived from the *alpha-fair utility model* [23], [17]. This model does not require fine-grained information about the location and behavior of users and therefore it is more suitable for large user populations than the MSD model. It only requires the estimation of the following values; the *price sensitivity* $r \in (1, +\infty)$ and the *valuation coefficient* v_n for every AP site n, i.e., $N+1$ values in total. Specifically, the demand for a WiFi AP n can be written as:

$$Q_n(\boldsymbol{x}, p_a) = \left(\frac{v_n}{p_a}\right)^r \mathbf{1}_{\{x_n=1\}} \ . \tag{4}$$

where the demand will be zero for non deployed APs ($x_n \neq 1$). For a given price p_a and sensitivity r, the demand will be higher for sites with larger v_n values, capturing the users' preferences among sites. For example, sites close to popular venues are expected to attract more user demand (and hence have higher v_n) than the rest. Higher values of r indicate higher elasticity; users are more sensitive to changes in price p_a. In other words, when r is high users respond to price changes in a more dramatic way. Figure 2 presents an example of the CED function for an AP with $v_1 = 5$, and three values of r, 1, 2 and 3.

The demand offloaded to the APs will be subtracted from the original cellular network demand. Hence, the latter is written:

$$Q_b(\boldsymbol{x}, p_a) = K - \sum_{n \in \mathcal{N}_{\boldsymbol{x}}} Q_n(\boldsymbol{x}, p_a) \ , \tag{5}$$

where, for simplicity, we have assumed that $\sum_{n \in \mathcal{N}_{\boldsymbol{x}}} Q_n(\boldsymbol{x}, p_a) \leq K$.

A feature of the CED model is that the demands of APs are *separable*, i.e., changes in the demand of an AP have no effect on the demand of the rest APs (cf. Eq. (4)). Thus, this model is appropriate when the *deployment of APs is sparse* and each user is in range with at most one AP. In dense AP deployments, however, the users will be in range with multiple APs meaning that there will be dependencies of the demand across them. To capture such dependencies, we will also study the following (more complicated, yet more accurate) demand model.

iii. Logit demand (LD). The LD model is common in the economic literature, and it has been recently used in Internet transit [17] and residential broadband [18] studies. The demand is *probabilistically* determined based on a set of *surplus* values associated to the AP sites. Namely, the surplus of the AP at site n is defined as the difference between the utility value v_n and the payment p_a. The probability that a user will be assigned to the AP at site n is given by:

$$s_n(\boldsymbol{x}, p_a) = \frac{e^{v_n - p_a}}{1 + \sum_{n' \in \mathcal{N}_{\boldsymbol{x}}} e^{v_{n'} - p_a}} \mathbf{1}_{\{x_n=1\}} \ , \tag{6}$$

i.e., the demand depends both on the surplus of the assigned AP and the other deployed APs. Figure 3 presents an example of the LD function for two APs where we fix the price p_a to 5 and the valuation coefficient v_1 to 5, but we vary v_2 between 0 and 10. We observe that increasing the value of v_2 causes more demand to be served by AP 2, while at the same time the demand of AP 1 is reduced. In other words, as AP 2 becomes more valuable for users, they shift their traffic from AP 1 to AP 2. These dynamics cannot be captured by the CED model, which assumes that the demands of the APs are separable.

The probability of assigning traffic to the cellular BSs will be:

$$s_b(\boldsymbol{x}, p_a) = 1 - \sum_{n \in \mathcal{N}_{\boldsymbol{x}}} s_n(\boldsymbol{x}, p_a) = \frac{1}{1 + \sum_{n \in \mathcal{N}_{\boldsymbol{x}}} e^{v_n - p_a}} \ . \tag{7}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2919590, IEEE Transactions on Communications

Therefore, the demand that will be offloaded to the AP at site n is given by:

$$Q_n(\boldsymbol{x}, p_a) = s_n(\boldsymbol{x}, p_a) \cdot K \ . \tag{8}$$

Similarly, $Q_b(\boldsymbol{x}, p_a) = s_b(\boldsymbol{x}, p_a) \cdot K$ denotes the user demand served by the cellular network.

LD requires the estimation of only N values (the v_n values), one for each possible AP site. Thus, similarly to CED, LD is more suitable for large user populations than the MSD model.

C. WiFi AP Deployment and Pricing Problem

For any of the demand models, the carrier profit will be the difference between revenue (payments) and costs:

$$\Pi(\boldsymbol{x}, p_a) = \sum_{n \in \mathcal{N}_{\boldsymbol{x}}} \Big((p_a - c_a) Q_n(\boldsymbol{x}, p_a) - d_n \Big) + (p_b - c_b) Q_b(\boldsymbol{x}, p_a) \ . \tag{9}$$

Since $c_a < c_b$, the *cost savings* increase with the volume of traffic offloaded to the WiFi APs. On the other hand, shifting traffic to APs causes a *revenue loss* when $p_a < p_b$. Therefore, the cost savings and revenue loss need to be balanced.

The Joint WiFi Deployment and Pricing (JWDP) problem asks for *jointly* determining the values in $\boldsymbol{x}$ and p_a in a way that maximizes the carrier's profit:[6]

$$max_{\boldsymbol{x}, p_a} \Pi(\boldsymbol{x}, p_a) \tag{10}$$
$$s.t. \ p_a \geq 0 \ , \tag{11}$$
$$x_n \in \{0, 1\}, \ \forall n \in \mathcal{N} \ . \tag{12}$$

The solution will reveal how a carrier should deploy APs and decide prices, and which parameters will affect the offloading benefits. This is a challenging optimization problem that includes both discrete ($\boldsymbol{x}$) and continuous (p_a) variables. Moreover, the objective function is in general non-linear due to the presence of the demand functions in Eq. (9).

III. AP DEPLOYMENT AND PRICING POLICIES

In this section, we address the JWDP problem for each of the demand models.

A. Approximate solution under MSD model

In this subsection, we present an approximate solution to the JWDP problem under the MSD model. By approximate we mean that there is a bound on the worst case performance of the solution. We first need to formulate the problem in a mathematical form that is amenable to efficient solution. To this end, we introduce the auxiliary variable y_{kn} which depends on p_a and $\boldsymbol{x}$ and indicates whether user k will offload her data to WiFi AP n ($y_{kn} = 1$) or not ($y_{kn} = 0$). Similarly, y_{kb} indicates the decision for the cellular network. The respective vector of variables is given by $\boldsymbol{y} = (y_{kn} \ : \ \forall k \in \mathcal{K}, \ n \in \mathcal{N} \cup \{b\})$. Then, we obtain the following formulation.

[6]We note that we optimize only p_a while treating p_b as constant and input to our problem. This is because, p_b typically depends on several different factors that are our of scope of this paper such as the competition with other carriers in the same market.

$$\max_{\boldsymbol{x}, p_a, \boldsymbol{y}} \sum_{n \in \mathcal{N}} \Big(\sum_{k \in \mathcal{K}} (p_a - c_a) y_{kn} - d_n x_n \Big) + \sum_{k \in \mathcal{K}} (p_b - c_b) y_{kb}$$

$$s.t. \ \text{constraints: } (11), \ (12)$$
$$y_{kn} = 0, \forall k \in \mathcal{K}, n \notin \mathcal{N}_k \tag{13}$$
$$y_{kn} \leq x_n, \forall k \in \mathcal{K}, n \in \mathcal{N}_k \tag{14}$$
$$y_{kn}(l_{ka} - p_a) \geq y_{kn}(l_{kb} - p_b), \forall k \in \mathcal{K}, n \in \mathcal{N}_k \tag{15}$$
$$\sum_{n \in \mathcal{N}_k \cup \{b\}} y_{kn} = 1, \forall k \in \mathcal{K} \tag{16}$$
$$y_{kn} \in \{0, 1\}, \forall k \in \mathcal{K}, n \in \mathcal{N} \cup \{b\} \tag{17}$$

Eq. (13) and (14) forbid users from offloading data to APs that are out of communication range or APs that have not been deployed. Eq. (15) ensures that a user will offload her data to an AP n only if this yields higher surplus than the cellular network. We emphasize that the offloading decision y_{kn} is made by the user not the carrier. Yet, the carrier can predict this decision based on the MSD model which regards the user as a self-interested entity who will pick the decision that maximizes her own surplus. With recent technology advances, the offloading decision can be coordinated by the carrier which can assign the user to the AP having the currently best available signal or the lowest load or based on other criteria (e.g., see Google's Project Fi service [6]). Finally, Eq. (16) states that each user will be served once.

To solve this problem the following lemma is needed. We defer the proofs to the Appendix that is available as a supplementary material of this paper.

Lemma 1: In the optimal solution of the JWDP problem under the MSD model it holds $p_a \in \{l_{ka} - (l_{kb} - p_b) : k \in \mathcal{K}\}$.

Lemma 1 states that the optimal price p_a belongs to a finite set of values. Therefore, we can solve the JWDP problem by computing for each of the K possibly optimal prices p_a the vectors $(\boldsymbol{x}, \ \boldsymbol{y})$ and then simply picking the price value that yields the highest carrier profit. We note that if we pick a price p_a for which the difference $p_a - c_a$ is lower than or equal to $p_b - c_b$, then the carrier will not be able to increase its profit by offloading traffic from the cellular network to the WiFi network. Hence, the optimal solution in this case will be straightforward; simply deploy no APs and offload no traffic ($x_n = 0, \ y_{kn} = 0 \ \forall n, k$), which yields carrier profit equal to $K(p_b - c_b)$. For any other possibly optimal price $p_a > p_b - c_b + c_a$, computing the optimal values of $(\boldsymbol{x}, \ \boldsymbol{y})$ is not straightforward. Interestingly, we can show that this subproblem is equivalent to the following facility location-type problem:

MAXUFLP [19]: We are given a set of potential facility locations $\mathcal{I}$ and a set of clients $\mathcal{J}$. Each client can be assigned to at most one facility. A cost o_i is associated with opening a facility at location $i \in \mathcal{I}$ and a revenue m_{ij} is associated with assigning client j to facility i. The question is to find a subset of facilities $\mathcal{S} \subseteq \mathcal{I}$ to open and assign clients to open facilities such that the total revenue minus cost is maximized.

Lemma 2: For a given pricing policy $p_a > p_b - c_b + c_a$, the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2919590, IEEE Transactions on Communications

Algorithm 1: Randomized Rounding for a given p_a

1 **if** $p_a \leq p_b - c_b + c_a$ **then**
2     $x_n = 0,\ y_{kn} = 0\ \forall n \in \mathcal{N},\ k \in \mathcal{K}$.
3 **else**
4     $\mathcal{K}' = \emptyset$.
5   **for** $k \in \mathcal{K}$ **do**
6   **if** $l_{ka} - p_a \geq l_{kb} - p_b$ **then**
7             $\mathcal{K}' = \mathcal{K}' \cup \{k\}$.
    end
8   Let JWDP$'$ be the JWDP problem given p_a where we replaced $\mathcal{K}$ for $\mathcal{K}'$.
9   Let $(\boldsymbol{x}^\dagger, \boldsymbol{y}^\dagger)$ be the optimal solution to the linear relaxation of JWDP$'$.
10   For each AP $n \in \mathcal{N}$, set $x_n = 1$ with probability $1 - 2(\sqrt{2}-1)(1 - x_n^\dagger)$, otherwise $x_n = 0$.
11   For each user $k \in \mathcal{K}'$, set $y_{kn} = 1$ for an AP $n \in \mathcal{N}_k$ with $x_n = 1$ if it yields higher surplus than the BSs, otherwise set $y_{kb} = 1$.
end
12 **return** $(\boldsymbol{x}, \boldsymbol{y})$

subproblem of computing the optimal values of $(\boldsymbol{x},\ \boldsymbol{y})$ under the MSD model can be reduced to the MAXUFLP.

The above lemma is important as it allows us to exploit the approximation algorithms that have been proposed for the MAXUFLP problem in order to solve our subproblem. A linear relaxation and randomized rounding based algorithm was shown to achieve an approximation ratio of $2(\sqrt{2}-1) \approx 0.828$ for MAXUFLP in [19]. The approximation measure used in this work is the ratio $\frac{f^{apx} - f^{min}}{f^{opt} - f^{min}}$ where f^{apx} is the objective value returned by the approximation algorithm, f^{opt} is the optimal objective value and f^{min} is a lower bound on the minimum objective value. In our problem, we can set $f^{min} = -\sum_{n \in \mathcal{N}} d_n < 0$, a case that appears when APs are deployed to all the sites and no revenue is earned. Hence, we obtain the following theorem.

Theorem 1: There is a solution to JWDP problem under the MSD model with profit f^{apx} such that $\frac{f^{apx} - f^{min}}{f^{opt} - f^{min}} \geq 2(\sqrt{2} - 1)$.

In the rest of this section, we describe the linear relaxation and randomized rounding-based algorithm in the context of our problem and present the pseudocode in Algorithm 1. This algorithm will be executed K times, one time for each possible optimal price p_a (cf. Lemma 1). In the end, we simply pick the price value that yields the highest carrier profit.

To start, for a given price p_a, we check whether shifting traffic from the cellular network to the WiFi APs can bring benefit to the carrier (line 1). If not, the optimal solution is simply to deploy no APs and offload no traffic (line 2). Otherwise (line 3), we define with JWDP$'$ the instance of the JWDP problem where we keep only the users whose surplus is higher for the WiFi than the cellular network, denoted by the $\mathcal{K}'$ set (lines 4-8). Then, we introduce the linear relaxation of the JWDP$'$ problem, which differs from JWDP$'$ in that the variables in $\boldsymbol{x}$ and $\boldsymbol{y}$ lie in $[0, 1]$, i.e., constraints (12) and (17) are replaced by $x_n \in [0, 1],\ \forall n \in \mathcal{N}$, and $y_{kn} \in [0, 1],\ \forall k \in \mathcal{K},\ n \in \mathcal{N} \cup \{b\}$. Having found an optimal (fractional) solution to this problem, denoted by $(\boldsymbol{x}^\dagger, \boldsymbol{y}^\dagger)$ (line 9), we independently deploy an AP at each location $n \in \mathcal{N}$ with probability $1 - 2(\sqrt{2}-1)(1 - x_n^\dagger)$ (line 10). Finally, a user in $\mathcal{K}'$ will be assigned to a deployed AP if it is in communication range and yields higher surplus than the BSs (line 11).

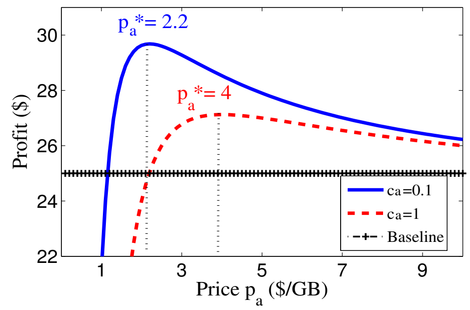


Fig. 4. Profit for a single AP under the CED model.

B. Optimal solution under CED model

We will show that the JWDP problem under the CED model can be optimally solved in polynomial time.

Theorem 2: The optimal AP deployment and pricing policy for the JWDP problem under the CED model can be found in polynomial-time.

Proof: By substituting Eq. (4) and (5) into (9), the carrier profit becomes:

$$\Pi(\boldsymbol{x}, p_a) = \sum_{n \in \mathcal{N}_{\boldsymbol{x}}} \left((p_a - c_a)\left(\frac{v_n}{p_a}\right)^r - d_n \right) + (p_b - c_b)\left(K - \sum_{n \in \mathcal{N}_{\boldsymbol{x}}} \left(\frac{v_n}{p_a}\right)^r \right). \quad (18)$$

Then, we can differentiate Eq. (18) with respect to p_a to find the profit-maximizing price:

$$p_a^* = \frac{r(p_b - c_b + c_a)}{r - 1}. \quad (19)$$

Figure 4 illustrates the carrier profit for a single AP and different prices p_a. We set $v_1 = 1$, $d_1 = 1$, $r = 2$, $p_b = 10$, $c_b = 9$ and $K = 25$. We compare three cases; the case when the AP is not deployed (Baseline), and two more cases that differ in the access cost c_a of the AP. We find that when $c_a = 0.1$ the profit-maximizing price p_a^* is 2.2, but increases to $p_a^* = 4$ when $c_a = 1$. This indicates that *as the technology improves the transmission capabilities and the cost falls, the operator can offer lower prices, and we can quantify the exact prices*. We also find that *lowering c_a makes profit more sensitive to the charged price, and hence the money loss due to possible changes in the system parameter values will be greater.*

By substituting p_a^* into Eq. (18), the profit becomes:

$$\Pi(\boldsymbol{x}, p_a^*) = = \left(\frac{r(p_b - c_b + c_a)}{r-1} - p_b + c_b - c_a\right)\left(\frac{r-1}{r(p_b - c_b + c_a)}\right)^r \sum_{n\in\mathcal{N}_{\boldsymbol{x}}} {v_n}^r - \sum_{n\in\mathcal{N}_{\boldsymbol{x}}} d_n + (p_b - c_b)K \ . \tag{20}$$

Here, deploying an AP at site n increases revenue by $\left(\frac{r(p_b-c_b+c_a)}{r-1} - p_b + c_b - c_a\right)\left(\frac{r-1}{r(p_b-c_b+c_a)}\right)^r {v_n}^r$ but incurs a deployment cost d_n. In order to maximize profit, an AP should be deployed at site n if and only if the increase in revenue is higher than the deployment cost. Therefore, the optimal AP deployment policy $\boldsymbol{x}^*$ will be:

$$x_n^* = \begin{cases} 1, & \text{if } v_n \geq \sqrt[r]{\dfrac{d_n}{\left(\frac{r(p_b-c_b+c_a)}{r-1} - p_b + c_b - c_a\right)\left(\frac{r-1}{r(p_b-c_b+c_a)}\right)^r}} \\ 0, & \text{otherwise} \end{cases} \tag{21}$$

■

Eq. (21) reveals how the deployment decisions depend on the system parameters. It is not difficult to see that the right-hand side of Eq. (21) increases monotonically with each one of the terms d_n, $p_b - c_b$, c_a and r. Therefore, we infer that *the optimal number of deployed APs is a decreasing function of the deployment cost (d_n), the net profit for serving traffic by the BS ($p_b - c_b$), the WiFi link cost (c_a) and the price sensitivity (r).*

Theorem 2 is important, since it states that the carrier can optimally deploy APs by simply evaluating a condition for each AP site. The overall complexity is $O(N)$. Nevertheless, as will show in the next subsection, the JWDP problem becomes challenging under the LD model.

C. Approximate solution under LD model

In this subsection, we present an approximate solution to the JWDP problem under the LD model. Our methodology is to express the problem as the maximization of a submodular set function. Before that, we need to describe the optimal pricing policy (p_a) for a given AP deployment policy ($\boldsymbol{x}$). Specifically, we prove the following lemma:

Lemma 3: For a given AP deployment $\boldsymbol{x}$, the optimal pricing policy under the LD model is given by:

$$p_a^* = w_{\boldsymbol{x}} + 1 + c_a + p_b - c_b \tag{22}$$

where

$$w_{\boldsymbol{x}} = \mathcal{W}\Big(\sum_{n\in\mathcal{N}_{\boldsymbol{x}}} e^{v_n - c_a - p_b + c_b - 1}\Big) \tag{23}$$

and $\mathcal{W}$ is the Lambert function, i.e., for any nonnegative z, $W(z)$ is the solution w satisfying $z = we^w$. Moreover, the optimal profit is given by:

$$\Pi^* = Kw_{\boldsymbol{x}} + K(p_b - c_b) - \sum_{n\in\mathcal{N}_{\boldsymbol{x}}} d_n \tag{24}$$

Figure 5 illustrates the carrier profit for different prices p_a under the LD model. We assume a setting with two APs and two values for the valuation v_n; 2 and 5. We further set $c_a = 1$,

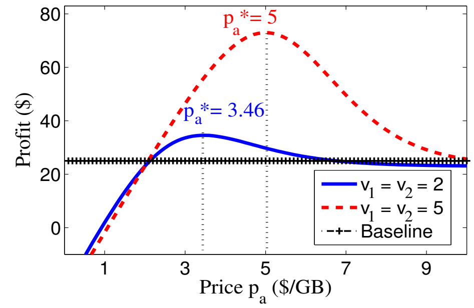


Fig. 5. Profit for two APs under the LD model.

$d_1 = d_2 = 1$, $p_b = 10$, $c_b = 9$ and $K = 25$. We observe that the profit and the optimal price increase with v_n. In the figure, the optimal price is $p_a^* = 3.46$ for $v_1 = v_2 = 2$, but increases to $p_a^* = 5$ for $v_1 = v_2 = 5$. In other words, as the users value more the APs, they are willing to offload more data at higher prices and this increases the carrier profit. Nevertheless, *making the p_a value very large can even result in lower profit compared to the case that none of the APs is deployed (Baseline)*. This is because the high price discourages most users from offloading data. Hence, the AP deployment costs (d_1 and d_2) can not be justified by the offloading cost savings.

Building on this lemma, we now present a class of approximation algorithms. We begin with the definition of submodular set functions:

Definition 1: Let G be a finite set of elements, which we refer to as ground set. A set function $f : 2^G \rightarrow \mathbb{R}$ is submodular if for all subsets $A, B \subseteq G$ with $A \subseteq B$ and every element $i \in G \setminus B$ we have:

$$f(A \cup \{i\}) - f(A) \geq f(B \cup \{i\}) - f(B) \tag{25}$$

In other words, the marginal value for adding an element i in a set decreases as the respective set expands. In our problem, we define the ground set G as follows:

$$G = (X_n : n \in \mathcal{N}) \tag{26}$$

where element X_n denotes the AP deployment at site n and it is the equivalent of deciding $x_n = 1$. A subset of elements $X \subseteq G$ corresponds to a deployment policy $\boldsymbol{x}$, such that $x_n = 1$ if and only if $X_n \in X$.

Let $\boldsymbol{x}_X$ be the binary representation of the set of elements X. Then, the objective in (10) can be written as a set function:

$$f(X) = max_{p_a \geq 0} \Pi(\boldsymbol{x}_X, p_a). \tag{27}$$

We can show the following lemma:

Lemma 4: The set function $f(X)$ is submodular.

Based on Lemma 4, the joint AP deployment and pricing problem is equivalent to maximizing a submodular set function f. There exist various approximation algorithms for this type of problems, especially when f is a monotone function. However, in our case, *f is not necessarily monotone*, since

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2919590, IEEE Transactions on Communications

Algorithm 2: Randomized Greedy

$A \leftarrow \emptyset, B \leftarrow G$
for $n = 1$ *to* N **do**
  $\Delta A \leftarrow g(A \cup \{u_n\}) - g(A)$
  $\Delta B \leftarrow g(B \setminus \{u_n\}) - g(B)$
  $\Delta A \leftarrow \max(\Delta A, 0)$, $\Delta B \leftarrow \max(\Delta B, 0)$
  with probability $\Delta A/(\Delta A + \Delta B)^*$ **do:**
  $A \leftarrow A \cup \{u_n\}$
  else (with probability$\Delta B/(\Delta A + \Delta B)$) **do:**
  $B \leftarrow B \setminus \{u_n\}$
end
return A (or equivalently B)

* If $\Delta A = \Delta B = 0$, then $\Delta A/(\Delta A + \Delta B) = 1$.

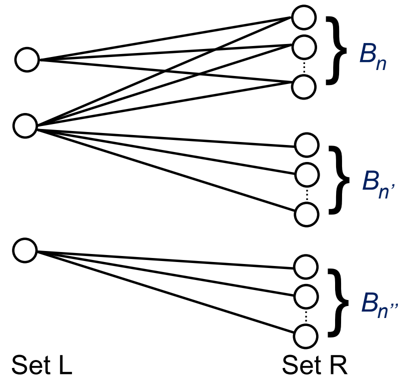


Fig. 6. Finding the optimal offloading policy $\boldsymbol{y}$ under the MSD model is equivalent to a maximum bipartite matching with node sets L and R. The node set L consists of the users whose surplus for WiFi is higher than cellular data (i.e., $l_{ka} - p_a \geq l_{kb} - p_b$). The node set R contains B_n nodes for each deployed AP (i.e., $x_n = 1$). Each user-node is connected to the AP-nodes that are in communication range (i.e., $n \in \mathcal{N}_k$).

adding an element to a set X increases $w_{\boldsymbol{x}_X}$, but at the same time yields a deployment cost (cf. Eq. (24)). The balance between these two factors determines whether the f value will increase or decrease. Even without the monotonicity property, there exist several algorithms with provable approximation guarantees for maximizing submodular functions requiring only the function f to be nonnegative. For example, a randomized greedy based algorithm was shown to achieve an $1/2$ approximation bound in [20]. This means that the ratio of the value of the approximate solution over the optimal solution value is always at least $1/2$, namely $f^{apx}/f^{opt} \geq 1/2$.

However, in our problem, f can take negative values. The lowest value is $f^{min} = -\sum_{n\in\mathcal{N}} d_n < 0$, a case that appears when APs are deployed to all sites and no data is offloaded. Hence, the above approximation bound does not hold. An idea is to maximize a normalized (non-negative) submodular function g where $g(X) = f(X) + f^{min}$ instead of f. Then, it is easy to show that the approximation bound holds for the expression $\frac{f^{apx}-f^{min}}{f^{opt}-f^{min}}$.

Theorem 3: There is a solution to JWDP problem under LD model with profit $\frac{f^{apx}-f^{min}}{f^{opt}-f^{min}} \geq \frac{1}{2}$.

As we summarize in Algorithm 2, the randomized greedy based algorithm proceeds in N iterations that correspond to some arbitrary order $u_1, \ldots u_N$ of the ground set G. At each iteration, it maintains two solutions A and B, which are initialized to $\emptyset$ and G respectively. At the n^{th} iteration, the algorithm either adds u_n to A or removes u_n from B. This decision is being made randomly based on the marginal gain of each of the two options. Eventually, after N iterations both solutions coincide, and we get $A = B$; this is the output of the algorithm.

IV. Impact of AP Congestion

So far, we employed a practical model that is accurate when the demand is lower than the available capacity, as in (sub) urban areas where the APs are supported by high-bandwidth backhaul lines[7]. This model can be easily augmented with additional components to account for practical constraints. For example, an AP may not be capable to serve all the user traffic in its neighborhood and have risk to become congested. To handle such cases, a bandwidth capacity bound B_n can be applied on the traffic offloaded to each AP $n \in \mathcal{N}$. This should be incorporated into the definition of the demand models and the problem formulation. Below, we describe how this can be accomplished for the MSD model.

Formally, the set of constraints in (11)-(17) is expanded to include the following bandwidth capacity constraint.

$$\sum_{k\in\mathcal{K}} y_{kn} \leq B_n, \ \forall n \in \mathcal{N} \tag{28}$$

The above constraint makes the JWDP problem more complex. On the positive side, Lemma 1 still holds as its proof depends only on constraint (15). Hence, as we explained in Subsection III-A, in order to solve it suffices to compute for each possibly optimal price p_a the vectors $(\boldsymbol{x}, \boldsymbol{y})$. This subproblem will be again a facility location type problem like MAXUFLP. The only difference is that each facility can serve up to a number of clients, corresponding to the capacity of the respective AP.

The capacitated facility location problem has been extensively studied in the literature (e.g., see [25]). Interestingly, recent studies have shown that this problem can be casted as the maximization of a submodular function [26], [27]. Therefore, it can be solved by running the same randomized greedy algorithm used for the LD model (Algorithm 2). The only difference is that $f(X)$ will be the carrier profit under the MSD model, rather than the one given by Eq. (24).

To understand this, note that for a given pricing p_a and AP deployment $\boldsymbol{x}$, the optimal offloading policy $\boldsymbol{y}$ can be computed in polynomial time by solving a bipartite matching problem, as it is depicted in Figure 6. Therefore, we can express the carrier profit under the MSD model as a function of the deployment policy only:

$$f(X) = max_{\{\boldsymbol{y}:(13)-(17),(28)\}} \Pi(\boldsymbol{x}_X, p_a, \boldsymbol{y}) \tag{29}$$

[7]A typical carrier-grade AP has several tens of Mbps of wireless capacity and is supported by a backhaul of similar capacity [24].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2919590, IEEE Transactions on Communications

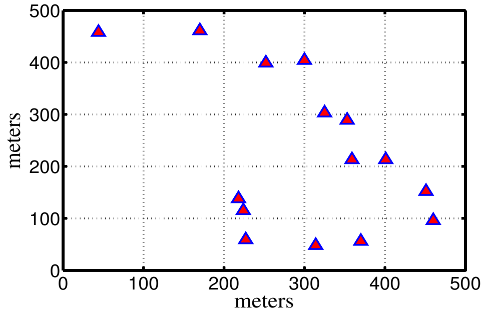


Fig. 7. A 500m×500m area with 15 WiFi APs [28].

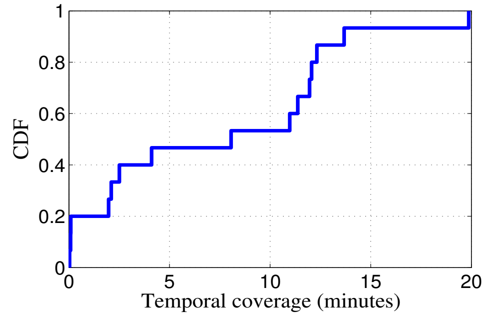


Fig. 8. The cumulative distribution function (CDF) of the temporal coverage in [28].

This can be shown to be a submodular function using the same arguments used in [26], [27]. Therefore, we obtain the following proposition.

Proposition 1: There is a solution to the capacitated version of the JWDP problem under the MSD model with profit $\frac{f^{apx}-f^{min}}{f^{opt}-f^{min}} \geq \frac{1}{2}$.

In the next section, we will evaluate the impact of AP capacities on the efficiency of the deployment and pricing decisions.

V. Dataset-driven Evaluation

We evaluate the impact of the proposed methods using a dataset of WiFi access patterns collected from real users. The evaluations cover a number of deployment cost and population density scenarios, which lead to an understanding of how the impact varies in different regions. We find that in mature wireless Internet markets deploying APs can be an effective mechanism for carriers to reduce their costs. Some virtual carriers can even halve their payments for using other carriers' cellular infrastructure, while charging users about 50% lower than the cellular service. However, deploying APs in a heuristic (ad-hoc) manner can actually lead to a net loss for the carrier, especially in less mature markets where the AP deployment or access costs are higher. In the rest of this section, we discuss these results in detail; we begin by describing the setup used in the later evaluations.

A. Evaluation setup

Since we do not know the valuation coefficient values l_{ka}, l_{kb} and v_n, which are introduced in Section II, we will extract them from the public WiFi access dataset in [28]. This dataset contains information from approximately 275 mobile users for an 11 week period. Specifically, each active user records every 20 seconds all the WiFi APs that are detected by her device. In total, more than 400 APs are detected. For each AP, its geographical location, described by a pair of (X,Y) values (measured in meters), is also recorded. For our evaluation, we consider a certain subarea in [28] with dense AP deployment depicted in Figure 7. To facilitate presentation, we shifted the point (X,Y)=(1698270,259950) to the zero coordinates. For each AP, we consider a candidate site for AP deployment, which yields $N = 15$ sites. We keep the data of the busiest day, namely the day of 16 October, and during peak-time, i.e., between 18pm and 23pm. We focus on this subset of the dataset for two reasons; (i) it represents a scenario with high user activity for which WiFi data offloading is needed most, and (ii) by limiting the number of AP sites we can apply exhaustive search to find the optimal solution in reasonable time. This is important, since it allows us to measure the optimality gap of the proposed algorithms.

We refer to the average time duration within which a user is covered by the WiFi AP at a site n in the dataset as the *temporal coverage* of this site, denoted by tc_n (minutes). Intuitively, the offloading benefits of deploying an AP at site n increase with tc_n, since users can offload data more often. Therefore, it is important to quantify this metric. Figure 8 shows the cumulative distribution function (CDF) of it for the considered 5-hour peak-time period. We observe a high disparity of the temporal coverage across the APs; for the less contacted AP tc_n is less than 1 minute, whereas on the upper extreme it is 20 minutes.

We will use the tc_n collected values to set up our evaluations. Since the dataset does not contain information about the data volumes of the users we will generate them artificially. Specifically, to model a realistic dense city scenario, we consider a population density equal to $12,000$ mobile users per square mile, as in [29]. This results in $K = 1158$ mobile users in the $500\text{m}\times500m$ area of the dataset. As a canonical scenario, we set the peak-time data traffic of each mobile user to 1GB/month [2]. We then randomly place the mobile users within the coverage regions of the APs following the tc_n values (and this way we form the respective $\mathcal{N}_k$ sets).

To find the valuation coefficient values of the demand models, we follow a similar method to [17]. First, we assume that the carrier charges $p_a = 5$ \$/GB and $p_b = 10$ \$/GB for downloading data from the WiFi and cellular network respectively. These prices are consistent with the current prices of small virtual carriers in the US market ([7], [8]). In response to the above prices, we assume that 80% of the traffic is directed to the WiFi network. This number is inline with recent reports about the WiFi traffic volume of virtual carriers [30]. To determine the volume of offloaded traffic by each AP n, we make the reasonable assumption that this is proportional

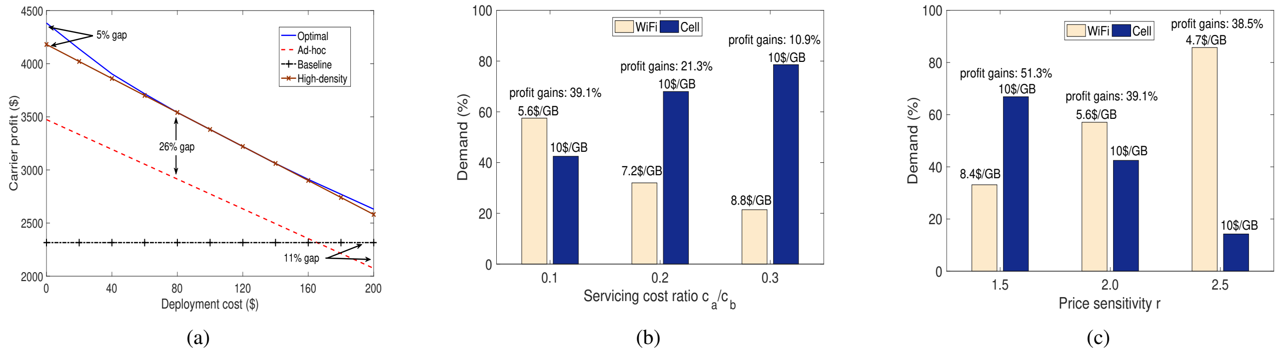


Fig. 9. Results for the CED model. (a) The impact of varying deployment cost on carrier profit. The distribution of demand over cellular and WiFi networks for (b) different servicing costs and (c) different price sensitivity parameters.

to its tc_n value. For the CED model we obtain (cf. Eq. (4)):

$$v_n^{ced} = (\frac{tc_n}{\sum_{n\in\mathcal{N}} tc_n} \cdot 0.80 \cdot 1158)^{\frac{1}{r}} \cdot p_a \ . \tag{30}$$

Unless otherwise specified, the price sensitivity parameter r is set to 2. Similarly, for the LD model we obtain (cf. Eq. (6)):

$$v_n^{ld} = p_a + \log(\frac{tc_n \cdot 0.80 \cdot 1158}{\sum_{n\in\mathcal{N}} tc_n}) - \log(0.20 \cdot 1158) \ . \tag{31}$$

Finally, for the MSD model, the l_{ka} value is picked randomly between \$5 and \$10 for 80% of the users, while for the rest 20% it is picked between \$0 and \$5. The l_{kb} is always 10 \$/GB.

The cellular network access cost c_b is set to 8.0\$/GB which is reasonable considering a 20% profit margin. The AP access cost c_a and the deployment costs d_n are varied throughout the evaluation. Unless otherwise specified, we set $c_a = 0.1c_b$ and $d_n = 120$\$/month (amortized) [29].

We will evaluate the following AP deployment and pricing algorithms:

1) *Baseline*: None WiFi AP is deployed. All users are served by the cellular network.
2) *Randomized-rounding*: Applies only to the MSD model. The joint AP deployment and pricing policy is found by running Algorithm 1, described in Section III-A.
3) *Randomized-greedy*: Applies only to the LD model and the capacitated case of the MSD model. The joint AP deployment and pricing policy is set by running Algorithm 2, described in Section III-C.
4) *Optimal*: The optimal solution to the JWDP problem. For the CED model, this is found based on the procedure in Theorem 1. For the rest models, Optimal is found through exhaustive search of the solution space. Since the required running time is exponential to the number of AP sites, Optimal is only used as a benchmark for gauging the performance of the proposed solutions and determining if there is still room for improvement.
5) *Ad-hoc*: This algorithm is used for comparison with the above proposed algorithms. The WiFi APs are deployed uniformly at random across the candidate sites, while the optimal price is selected.
6) *High-density*: This algorithm deploys the same number of WiFi APs with the Ad-hoc algorithm but picks the sites with the highest user demand instead of random sites. The optimal price is selected.

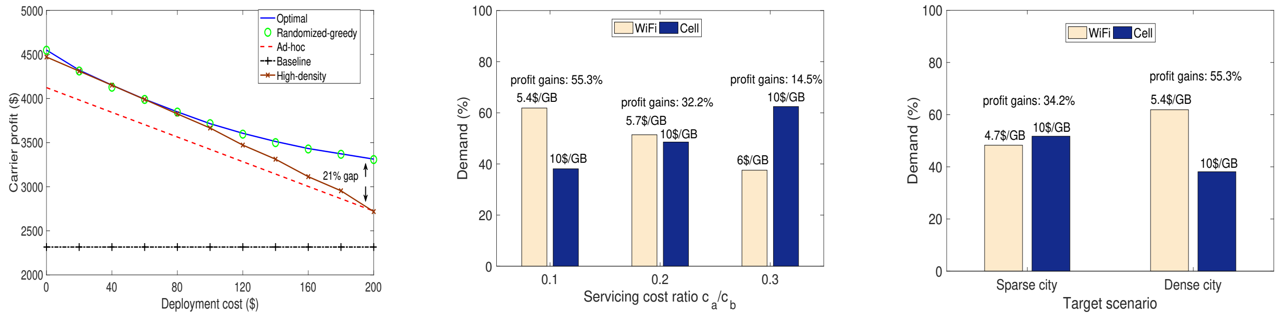
We remark that our evaluation code is publicly available online in [31] to facilitate future experimentation with WiFi data offloading solutions.

B. Evaluation results

We will evaluate the carrier profit achieved by the above algorithms under different demand models, while varying the AP deployment and servicing costs, the population density and the price sensitivity of users. The results will be presented separately for each demand model. We begin with the CED model.

Results for the CED model: Figure 9(a) depicts the carrier profit for different AP deployment costs $d_n \in [0, 200]$ \$/month $\forall n$. We note that these d_n values are amortized and inline with recent reports [29]. Higher d_n values typically refer to less mature markets where the absence of extensive wireline infrastructure may result in higher costs for the backhaul AP deployment. As expected, increasing the deployment costs reduces the carrier profit for all the algorithms, but the Baseline. *In mature markets (with low d_n values), deploying APs via the Optimal algorithm can even double the carrier profit compared to Baseline*. Deploying APs in an ad-hoc manner may yield up to 26% lower carrier profit than Optimal. In less mature markets, ad-hoc AP deployment can even lead to a net loss for the carrier (up to 11% lower profit compared to Baseline). Deploying APs to the sites with high user density achieves close to optimal carrier profit (up to 5% gap in the considered scenario), showing that this can be a practical deployment solution in certain cases. We need to emphasize, however, that for the evaluation of the High-density and Ad-hoc schemes we used the optimal price p_a uncovered by our analysis (equation (19)). The performance of these schemes will worsen arbitrarily if suboptimal prices are picked.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2919590, IEEE Transactions on Communications



(a) (b) (c)

Fig. 10. Results for the LD model. (a) The impact of varying deployment cost on carrier profit. The distribution of demand over cellular and WiFi networks for (b) different servicing costs and (c) different population density scenarios.

We next explore how the ratio c_a/c_b of servicing costs impacts the results (Figure 9(b)). Here, we keep $c_b = 8$ \$/GB constant, but we vary c_a within $\{0.1c_b, 0.2c_b, 0.3c_b\}$. The vertical axis depicts the percentage of user data demand served by the WiFi and cellular network in each case, corresponding to Optimal algorithm. We observe that *increasing the c_a cost reduces the percentage of demand served by the WiFi APs, e.g., from 57.5% when $c_a = 0.1c_b$ to 21.4% when $c_a = 0.3c_b$.* This can be explained by the fact that the carrier increases the p_a price by 3.2\$/GB (from 5.6 to 8.8 \$/GB) in response to the increase in c_a cost (cf. Eq. (19)). This will discourage some users from offloading mobile data, which in turn causes a reduction in the profit gains over Baseline (from 39.1% to 10.9%).

The CED model we use relies on the price sensitivity parameter (r). Therefore, it is important to analyze how the results are affected by varying this parameter. Towards this goal, we repeat the experiments for $r \in \{1.5, 2.0, 2.5\}$ (Figure 9(c)). We observe that the optimal price p_a^* decreases with r (cf. Eq. (19)), ranging from 8.4\$/GB to 4.7\$/GB, less than half of the cellular network price. In other words, *the more sensitive the users are to changes in price, the lower the p_a^* price and the higher the portion of demand served by the APs become*.

Results for the LD model: We now repeat the experiments for the LD model. We find that the curve of carrier profit versus deployment costs is similar to the CED model (Figure 10(a)). A difference is that the benefits of the WiFi deployment algorithms over Baseline are slightly higher in absolute value for the LD model. *The proposed algorithm (Randomized-greedy) operates very close to Optimal (less than 1% gap) and significantly better than Baseline (up to 96% gap). The gap from the Ad-hoc and High-density schemes is also significant (up to 21% for $d_n = 200$), showing the merits of the proposed optimization approach.* High-density performs better than Ad-hoc for low deployment costs, yet these benefits vanish as deployment costs increase.

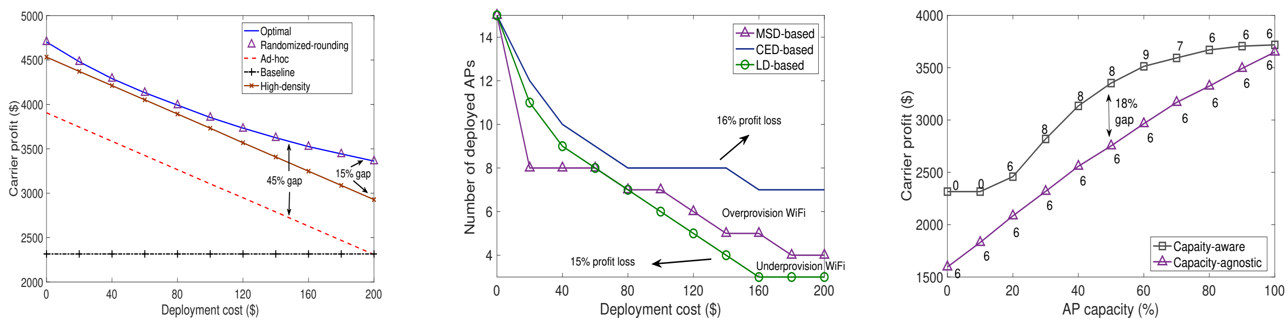
Similar results about the impact of servicing costs are observed in Figure 10(b). Again, the increase in WiFi servicing cost will result in an increase in price p_a. This increase in p_a is smoother (from 5.4\$/GB to 6\$/GB) than was under the CED model. Therefore, the respective change in demand is smoother as well.

The numerical results presented so far focused on a dense city scenario with high population density. To analyze how the results vary in different regions, we synthesize a second scenario of a sparse city with population density equal to $5,000$ (instead of $12,000$) people per square mile [29]. Figure 10(c) shows the distribution of demand over cellular and WiFi networks achieved by Randomized-greedy algorithm in each scenario. We observe that *the profit gains over Baseline are higher in the dense city than in the sparse, since the need for offloading data in the former is greater.* This observation is consistent with the fact that offloading mechanisms have existed for a long time in densely populated areas. In particular, the gains increase from 34.2% (sparse city) to 55.3% (dense city).

Results for the MSD model: Figure 11(a) shows the carrier profit versus AP deployment cost curve for the MSD model. The proposed Randomized-rounding algorithm achieves an optimal or near-optimal solution for all the deployment costs (less than 1% gap). Besides, it performs significantly better than Baseline, Ad-hoc and High-density algorithms. The profit gains are up to 45% for the Baseline and Ad-Hoc algorithms, and up to 15% for the High-density algorithm.

At this point, we emphasize that the three considered demand models take different information as input and, hence, the respective proposed algorithms make different decisions. Figure 11(b) shows the number of APs deployed for each demand model. We observe that the CED-based algorithm typically deploys more APs than the MSD-based algorithm, while the LD-based algorithm deploys fewer APs. This is an interesting result since it reveals that *depending on the demand model used, the solution tends to overprovision or underprovision the WiFi network.* Among the three demand models, MSD requires the most fine-grained information as input. The lack of complete information by the CED and LD models results in up to 16% and 15% lower profit compared

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2919590, IEEE Transactions on Communications



(a) (b) (c)

Fig. 11. Results for the MSD model. (a) Carrier profit for different deployment costs. (b) The number of deployed APs for different demand models. (c) The impact of AP capacity (as a percentage of the total data demand in the network) on carrier profit.

to MSD in the above experiment.

Finally, we investigate the impact of AP capacities on the efficiency of AP deployment and pricing decisions. Figure 11(c) compares the carrier profit achieved by the two approximation algorithms we proposed for the MSD model; one that neglects capacity constraints (capacity-agnostic) and one that considers these constraints (capacity-aware). The labels along the curves indicate the number of deployed APs in each case. When the capacities are very low ($< 20\%$ of the total demand), the deployment cost cannot be justified by the offloading benefits, and hence none AP is deployed by the capacity-aware algorithm. The number of deployed APs increases with the capacities (from 0 to 9) up to a point where it starts to decrease (from 9 to 6). *Neglecting AP capacities can result in underprovisioning or overpsosioning the WiFi network, which can cause up to* 18% *profit loss*.

VI. Related Work

Existing *measurement studies* [32] have shown that in urban areas WiFi can offload about 65% of mobile data, and have quantified the performance benefits of this approach. Despite this momentum, the AP deployment issue has received little attention thus far. There are only few systematic studies for sophisticated AP deployment. For example, the works [33], [34], [35], [36] and [37] analyzed such network design problems and proposed meaningful solutions. Nevertheless, these works do not consider the economic aspects of WiFi data offloading neglecting that the consensus of the end-users is required to offload their data.

Recently, several works have studied the economic aspects of WiFi data offloading. Schemes that lease third-party owned WiFi APs have been proposed in [38], [39], [40], but again neglecting end-user consensus. The works in [41] and [42] proposed monetary compensations for the end-users so as to delay their transmissions until they reach a WiFi AP, and hence increase the offloading benefits for the operators. A queueing analytic model for the performance of delayed mobile data offloading has been proposed in [43]. The economic interactions among the WiFi AP owners, advertisers and end-users, when the latter watch advertisements in exchange of the free usage of WiFi, have been studied in [44]. The economic consequences of opportunistic WiFi data offloading for both carriers and end-users have been analyzed in [5]. Nevertheless, these works do not consider jointly all the aspects that affect the end-users' decisions, and most importantly do not correlate them with the AP deployment strategies.

VII. Conclusion

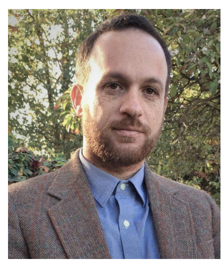
This paper introduced a new methodology for designing carrier-grade WiFi deployment and pricing policies. Despite the importance and momentum of WiFi, the deployment issue has received little attention by previous works. We addressed this problem for general and diverse models of user demand using both off-the-shelf (e.g., facility-location algorithms appropriately tailored to our problem) and new optimization approaches. Some of the presented results might be of independent interest. For example, the proof of submodularity under the LD model may be exploited for other network design and economic problems. Dataset-driven evaluations demonstrated that deploying WiFi can be an effective mechanism for carriers to reduce their costs. Some virtual carriers can even halve their payments for using other carriers' cellular infrastructure, while charging users about 50% lower than the cellular service.

References

[1] K. Poularakis, G. Iosifidis, L. Tassiulas, "Deploying Carrier-grade WiFi: Offload Traffic, Not Money", *ACM MobiHoc*, 2016.
[2] Ericsson, "Mobility Report", November 2018.
[3] A. Argyriou, K. Poularakis, G. Iosifidis, L. Tassiulas, "Video Delivery in Dense 5G Cellular Networks", *IEEE Network, vol. 31, no. 4, pp. 28-34*, 2017.
[4] Proceedings of the 1st IEEE 5G Summit, [Online]: http:www.5gsummit.org/index.htm, 2015.
[5] L. Zheng, J. Chen, C. Joe-Wong, C. W. Tan, M. Chiang, "An Economic Analysis of Wireless Network Infrastructure Sharing", *IEEE WiOpt*, 2017.
[6] Google, "Project Fi", fi.google.com, 2017.
[7] Republic Wireless, [Online]: www.wirefly.com/wireless-plan/republic-wireless-5-plan-wifi-only
[8] Republic Wireless Google Chromcast, [Online]: www.eightymphmom.com/2014/08/republic-wireless.html
[9] "US Mobile Announces $10 WiFi Only Plan With Worldwide Availability Including In Flight", [Online]: https://bestmvno.com/us-mobile/us-mobile-announces-10-wifi-plan-worldwide-availability-including-flight
[10] Boingo, [Online]: http://www.boingo.com/retail

[11] Comcast, Xfinity WiFi, [Online]: https://wifiondemand.xfinity.com/wod/#offerList
[12] FCC, "FCC Increases 5GHz Spectrum for Wi-Fi, Other Unlicensed Uses", 2014.
[13] 3GPP, "Technical Specification Group Services and System Aspects; Architecture enhancements for non-3GPP accesses (Release 13)", *3GPP TS 23.402 V13.3.0 (2015-09)*.
[14] Next Generation Mobile Networks (NGMN) Alliance, "5G White Paper", 2015, [Online]: www.ngmn.org
[15] AT&T, "AT&T Wi-Fi small site", Product Brochure, Dallas, TX, USA, Mar. 2015
[16] Facebook, [Online]: https://connectivity.fb.com/wi-fi
[17] V. Valancius, C. Lumezanu, N. Feamster, R. Johari, V.V. Vazirani, "How many tiers?: Pricing in the Internet Transit Market", *ACM Sigcomm*, 2011.
[18] K. Poularakis, I. Pefkianakis, J. Chandrashekar, L. Tassiulas, "Pricing the Last Mile: Data Capping for Residential Broadband", *ACM CoNEXT*, 2014.
[19] A.A. Ageev, M.I. Sviridenko, "An 0.828-approximation Algorithm for the Uncapacitated Facility Location Problem", *Discrete Applied Mathematics, vol. 93, pp. 149-156*, 1999.
[20] N. Buchbinder, M. Feldman, J. Naor, R. Schwartz, "A Tight Linear Time (1/2)-Approximation for Unconstrained Submodular Maximization", *IEEE FOCS*, 2012.
[21] J. Blackburn, R. Stanojevic, V. Erramilli, A. Iamnitchi, K. Papagiannaki. "Last Call for the Buffet: Economics of Cellular Networks", *ACM MobiCom*, 2013.
[22] R. Shioda, L. Tunçel, T.G.J. Myklebust, "Maximum Utility Product Pricing Models and Algorithms based on Reservation Prices", *Computational Optimization and Applications vol. 48, pp. 157-198*, 2011.
[23] J. Mo and J. Walrand, "Fair End-to-end Window-based Congestion Control", *IEEE/ACM Transactions on Networking, vol. 8, no. 5, pp. 556-567*, 2000.
[24] I. Fogg, "The State of Wifi vs Mobile Network Experience as 5G Arrives", *opensignal.com*, November 2018.
[25] F.A. Chudak, D. P. Williamson, "Improved Approximation Algorithms for Capacitated Facility Location Problems", *Mathematical Programming, vol. 102, no. 2, pp. 207-222*, 2005.
[26] E.M. Craparo, J.P. How, E. Modiano, "Throughput Optimization in Mobile Backbone Networks", *IEEE Transactions on Mobile Computing, vol. 10, no. 4, pp. 560-572*, 2011.
[27] T. Lukovszki, M. Rost, S. Schmid, "Approximate and Incremental Network Function Placement", *Journal of Parallel and Distributed Computing, vol. 120, pp. 159-169*, 2018.
[28] M. McNett, G.M. Voelker, "Access and Mobility of Wireless PDA Users", *ACM Sigmobile Mobile Computing and Communications Review, vol. 9, no. 2, pp. 40-55*, 2005.
[29] C. Joe-Wong, S. Seny, S. Ha, "Offering Supplementary Network Technologies: Adoption Behavior and Offloading Benefits", *IEEE/ACM Transactions on Networking, vol. 23, no. 2, pp. 355-368*, 2015.
[30] gigaom.com, [Online]: "Who's Your New Mobile Carrier? How 'bout Wi-Fi?", 2013.
[31] Publicly Available Online Code, https://www.dropbox.com/s/wrjr5kb81bwrepb/mobihoc16.rar?dl=0
[32] K. Lee, J. Lee, Y. Yi, I. Rhee, S. Chong, "Mobile Data Offloading: How Much Can WiFi Deliver?", *IEEE/ACM Transactions on Networking, vol. 21, no. 2, pp. 536-551*, 2013.
[33] C. S. Wang, L. F. Kao, "The Optimal Deployment of Wi-Fi Wireless Access Points Using the Genetic Algorithm", *ICGEC*, 2012.
[34] R. Prasad, and H. Wu, "Minimum-Cost Gateway Deployment in Cellular Wi-Fi Networks", *IEEE CCNC*, 2006.
[35] T. Wang, W. Jia, G. Xing, and M. Li, "Exploiting Statistical Mobility Models for Efficient Wi-Fi Deployment", *IEEE Transactions on Vehicular Technology, vol. 62, no. 1, pp. 360-373*, 2013.
[36] E. Bulut, and B. K. Szymanski, "Wi-Fi Access Point Deployment for Efficient Mobile Data Offloading", *ACM CCR, vol. 17, no. 1*, 2013.
[37] Z. Zheng, P. Sinha, S. Kumar, "Sparse WiFi Deployment for Vehicular Internet Access with Bounded Interconnection Gap", *IEEE/ACM Transactions on Networking, vol. 20, no. 3, pp. 956-969*, 2012.
[38] G. Iosifidis, L. Gao, J. Huang, L. Tassiulas, "A Double Auction Mechanism for Mobile Data Offloading Markets", *IEEE/ACM Transactions on Networking, vol. 23, no. 5, pp. 1634-1647*, 2015.
[39] H. Yu, M.H. Cheung, J. Huang, "Cooperative Wi-Fi Deployment: A One-to-Many Bargaining Framework", *IEEE Transactions on Mobile Computing, vol. 16, no. 6, pp. 1559-1572*, 2017.
[40] K. Poularakis, G. Iosifidis, I. Pefkianakis, L. Tassiulas, Martin May, "Mobile Data Offloading through Caching in Residential 802.11 Wireless Networks", *IEEE Transactions on Network and Service Management, vol. 13, no. 1, pp. 71-84*, 2016.
[41] J. Lee, Y. Yi, S. Chong, Y. Jim, "Economics of Wi-Fi Offloading: Trading Delay for Cellular Capacity", *IEEE Infocom Workshops*, 2013.
[42] F. Hsu, H. Su, "When Does the AP Deployment Incentivize a User to Offload Cellular Data: An Energy Efficiency Viewpoint", *IEEE ISCCSP*, 2014.
[43] F. Mehmeti, T. Spyropoulos, "Performance Modeling, Analysis, and Optimization of Delayed Mobile Data Offloading for Mobile Users", *IEEE/ACM Transactions on Networking, vol. 25, no. 1, pp. 550-564*, 2016.
[44] H. Yu, M.H. Cheung, L. Gao, J. Huang, "Public Wi-Fi Monetization via Advertising", *IEEE/ACM Transactions on Networking, vol. 25, no. 4, pp. 2110-2121*, 2017.
[45] H. Li, W. T. Huh, "Pricing Multiple Products with the Multinomial Logit and Nested Models: Concavity and Implications", Manufacturing & Service Operations Management, 2007.

Konstantinos Poularakis obtained the Diploma, and the M.S. and Ph.D. degrees in Electrical Engineering from the University of Thessaly, Greece, in 2011, 2013 and 2016 respectively. In 2014, he was a Research Intern with Technicolor Research, Paris. Currently, he is a Post-Doctoral researcher at Yale University. His research interests lie in the area of network optimization with emphasis on emerging architectures such as software defined networks and wireless caching networks. He was the recipient of several awards and scholarships during his studies, from sources including the Greek State Scholarships foundation, the Center for Research and Technology Hellas, the Alexander S. Onassis Public Benefit Foundation and the Bodossaki Foundation. He also received the Best Paper Award at the IEEE INFOCOM 2017 and the IEEE ICC 2019.



George Iosifidis received the Diploma degree in telecommunications engineering from the Greek Air Force Academy in 2000, and the M.S. and Ph.D. degrees in electrical engineering from the University of Thessaly in 2007 and 2012, respectively. He was a Post-Doctoral Researcher with CERTH, Greece, and Yale University, USA. He is currently the Ussher Assistant Professor in Future Networks with Trinity College Dublin, Ireland. His research interests lie in the broad area of wireless networks and network economics.

Leandros Tassiulas (S'89, M'91, SM/05 F/07) is the John C. Malone Professor of Electrical Engineering at Yale University. His research interests are in the field of computer and communication networks with emphasis on fundamental mathematical models and algorithms of complex networks, architectures and protocols of wireless systems, sensor networks, novel internet architectures and experimental platforms for network research. His most notable contributions include the max-weight scheduling algorithm and the back-pressure network control policy, opportunistic scheduling in wireless, the maximum lifetime approach for wireless network energy management, and the consideration of joint access control and antenna transmission management in multiple antenna wireless systems. Dr. Tassiulas is a Fellow of IEEE (2007). His research has been recognized by several awards including the IEEE Koji Kobayashi computer and communications award, the inaugural INFOCOM 2007 Achievement Award for fundamental contributions to resource allocation in communication networks, the INFOCOM 1994 and 2017 best paper awards, a National Science Foundation (NSF) Research Initiation Award (1992), an NSF CAREER Award (1995), an Office of Naval Research Young Investigator Award (1997) and a Bodossaki Foundation award (1999). He holds a Ph.D. in Electrical Engineering from the University of Maryland, College Park (1991). He has held faculty positions at Polytechnic University, New York, University of Maryland, College Park, and University of Thessaly, Greece.