LAPRA: Location-aware Proactive Resource Allocation

This document has been downloaded from Chalmers Publication Library (CPL). It is the author´s version of a work that was accepted for publication in:

IEEE Globecom Conference 2016

Citation for the published paper:

Downloaded from: http://publications.lib.chalmers.se/publication/245687

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.

(article starts on next page)
LAPRA: Location-aware Proactive Resource Allocation

L. Srikar Muppirisetty∗, Simon Yiu†, and Henk Wymeersch∗
∗Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden
†Bell Laboratories, Murray Hill, NJ, USA
E-mail: {srikar.muppirisetty, henkw}@chalmers.se, simon.yiu@nokia.com

Abstract—Today’s indoor wireless networks employ reactive resource allocation methods to provide fair and efficient usage of the communication system. However, their reactive nature limits the quality of service (QoS) that can be offered to the user locations within the environment. In large crowded areas (airports, conferences), networks can get congested and users may suffer from poor QoS. To mitigate this, we propose and evaluate a location-aware user-centric proactive resource allocation approach (LAPRA), in which the users are proactive and seek good channel quality by moving to locations where the signal quality is good. As a result, the users and their locations are optimized to improve the overall QoS. We demonstrate that the proposed proactive approach enhances the user QoS and improves network throughput of the system.

I. INTRODUCTION

Broadband service providers are faced with the daunting task of providing a consistent and seamless user experience [1], in the presence of three confluent factors. First of all, rising data demands, in particular streaming video is putting stress on both the core and access networks. Second, high variability in demands and user density lead to traffic surges, which operators must provision for, through network densification and intelligent resource allocation. Finally, the business models are such that operators cannot simply increase the price per byte consumed, and thus must deliver high quality of service (QoS), with minimal additional operational or capital expenses.

User QoS can be improved by means of either extending the existing infrastructure (e.g., adding more access points (APs), or/and employing smart resource allocation strategies. We can classify resource allocation as either reactive or proactive, depending on whether they are able to harness the user’s future characteristics, such as demand patterns, behavior, and channel characteristics. Predictability of demand patterns is considered in the context of proactive caching in [2]. In [3], a multi-user rate allocation method is proposed based on predicted user rates for efficient energy transmission of stored videos that can be cached at the user devices. Proactive caching based on user demand, and channel prediction, statistics is studied in [4]. A different approach, called user-in-the-loop (UIL), was proposed in [5], [6], where the user can be controlled, in either space or time. In spatial control of UIL, the users are directed to locations which increases the spectral efficiency of the network [5], [7]. A similar approach is considered in [8], where a user-centric load balance scheme for small cell networks is proposed, in which users are allowed to move and chose their serving APs, based on a simple path loss model. However, these works do not account for the uncertainties in real spatial channel prediction [9]–[11].

In this paper, we propose a location-aware proactive resource allocation approach (LAPRA) that accounts for uncertainties in spatial channel predictions. In particular, we evaluate a spatial UIL resource allocation method that utilizes Gaussian processes [12], [13] for channel prediction by exploiting user location information. We formulate a decentralized binary integer optimization problem for the user assignment. Each user decides where to move by solving an optimization problem. We analyze the performance of the proposed method, compared to a network-centric reactive approach and baseline network-centric proactive approach. We show that our proposed method improves the number of satisfied users and the overall network throughput.

II. MODEL AND BENCHMARK APPROACH

A. System Model

Consider a geographical region \( A \subset \mathbb{R}^2 \) consisting of a set of APs \( M = \{1, 2, \ldots, m, \ldots, M\} \) serving a set of users \( N = \{1, 2, \ldots, n, \ldots, N\} \). The location of the AP, \( m \in M \), is denoted as \( x_{m}^{AP} \) and the location of the user \( n \in N \) is denoted as \( x_{n} \). Let \( d_{n} \) be the required rate requirement of the user \( n \) and \( L_{\text{max}}^{m} \) be the maximum number of channels AP \( m \) can support. We assume rate demands \( d_{n} \) of the users are fixed. In order to serve more users, we assume each user will get at maximum one channel per time slot. We assume all the APs transmit with a fixed identical transmit power \( P_{TX} \). Based on standard wireless propagation models, the received power from the \( m \)-th AP to the \( n \)-th user can be expressed in dB scale as [14, Chap. 2]

\[
P_{RX}(x_{n})[\text{dBm}] = L_{0} - 10 \eta \log_{10}(\|x_{m}^{AP} - x_{n}\|) + \Psi(x_{n}),
\]

(1)

where \( L_{0} = P_{TX}[\text{dBm}] + 10 \log_{10}(g_{0}) \) and \( g_{0} \) is a constant that captures antenna and other propagation gains, \( \eta \) is the path-loss exponent, \( \Psi(x_{n}) \) is the location-dependent shadowing in dB scale. Shadowing \( \Psi(x_{n}) \) is spatially correlated, with well-established correlation models [15], among which the Gudmundson model is widely used [16]. In (1), we assume small-scale fading is averaged out in the measurements either
in time, frequency, or space. The supported rate, expressed in bits per channel use (bpu), by AP \( m \) to user \( n \) is defined as
\[
{r^m}(x_n) = \log_2(1 + SNR^m(x_n)),
\]
(2)
where \( SNR^m(x_n) \) is the signal-to-noise ratio between \( (m, n) \) AP-user pair. \( SNR^m(x_n) = \frac{P^m_{RX,lin}(x_n)}{\sigma_n^2} \), where \( \sigma_n^2 \) is the noise power and \( P^m_{RX,lin}(x_n) \) is the received power in linear scale.

We define the utility function for user \( n \) for the AP \( m \) as \( U_m^m(x_n) \). The utility function \( U_m^m(x_n) \) should be monotonic in \( SNR^m(x_n) \). As a working example, we will consider \( U_m^m(x_n) = r^m(x_n) \). Our objective is to assign users to APs in such a way that the overall system utility is met.

### B. Reactive Approach

In the reactive approach, the network simply performs user to AP associations. In particular, let \( y_{nm} \in \{0, 1\} \) be the variable which takes the value of 1 if \( n \)-th user is assigned to \( m \)-th AP and 0 otherwise. Then the reactive user assignment can be written as a binary integer optimization problem (BIP)

\[
\text{maximize} \quad \sum_{n=1}^N \sum_{m=1}^M U_m^m(x_n) y_{nm} \tag{3a}
\]

subject to
\[
\sum_{n=1}^N y_{nm} \leq L_m^m, \quad m \in M, \tag{3b}
\]
\[
d_n y_{nm} \leq r^m(x_n), \quad \forall n \in N, \quad \forall m \in M \tag{3c}
\]
\[
\sum_{m=1}^M y_{nm} \leq 1, \quad n \in N, \tag{3d}
\]
\[
y_{nm} \in \{0, 1\}, \quad n \in N, \quad m \in M. \tag{3e}
\]

We note the following: (3a) aims to maximize the rate of each user with respect to (w.r.t) to its rate requirement; (3b) makes sure the total number of channels served by the AP is less than the maximum supported channels of the corresponding AP; (3c) states that each user desired rate requirement should be met; (3d) guarantees that each user is assigned at most to one and only one AP; and (3e) imposes the binary integer requirements on the optimization variable.

### III. Proactive Approach

A problem with the reactive approach is that with an increase in the number of users, not all users’ rate requirements can be met. One way to deal with this is to allow the users to move to a different location. In order for users to determine better locations, they must have an estimate of the received power at those locations. In this section, we first describe a method where users can predict the SNR, and then outline the proposed proactive approach.

#### A. Channel Prediction Framework

In the proactive approach, the users can move to locations where the channel conditions are better. This requires a channel prediction engine which can provide the received power at each and every location of the region \( A \). We model the \( P^m_{RX}(x_n) \) as a Gaussian processes \([12],[17]\) with mean function \( \mu(x_{AP}^m,x_n) : A \rightarrow \mathbb{R} \) and a positive semi-definite covariance function \( C(x_{AP}^m,x_n) : A \times A \rightarrow \mathbb{R}_+ \)

\[
P^m_{RX}(x_n) \sim GP(\mu(x_{AP}^m,x_n),C(x_n,x_n')) \tag{4}
\]

The mean function is defined as \( \mu(x_{AP}^m,x_n) = \mathbb{E}_{\psi(x_n)}[P^m_{RX}(x_n)] = L_0 - 10 \eta \log_{10}(||x_{AP}^m - x_n||) \). The covariance function is defined as \( C(x_n,x_n') = \text{Cov}(P^m_{RX}(x_n),P^m_{RX}(x_n')) \). The Gudmundsson model \([16]\) is the most commonly used covariance function for shadowing, which is expressed as

\[
C(x_n,x_n') = \sigma^2 \exp \left(-\frac{||x_n-x_n'||}{d_c}\right) + \delta_{nn'} \sigma^2_{\text{mean}}, \tag{5}
\]

where \( \delta_{nn'} = 1 \) for \( n = n' \) and zero otherwise, \( d_c \) is the correlation distance of the shadowing, \( \sigma_{\text{mean}} \) captures the any residual un-modeled uncertainties. Note that both the mean function and the covariance depend on parameters \( \theta = [\sigma_{\text{mean}},d_c,L_0,\eta,\sigma_{\psi}] \).

**Learning:** The parameters \( \theta^m \) w.r.t. the AP \( m \) can be learned given the measurements at \( Q \) training locations \( X = [x_1^T, x_2^T, \ldots, x_Q^T]^T \) and the corresponding received power measurements \( y^m = [y^m(x_1),y^m(x_2),\ldots,y^m(x_Q)]^T \), \( \forall m \in M \). The model parameters can be learned through by minimizing the negative log-likelihood function with respect to \( \theta^m \):

\[
\theta^m = \arg \min_{\theta^m} \{-\log(p(y^m|X,\theta^m))\}. \tag{6}
\]

If we assume the underlying channel parameters does not change much w.r.t. to each AP, we can learn them only once.

**Prediction:** Once \( \hat{\theta}^m \) is obtained, we can determine the predictive distribution of \( P^m_{RX}(x^*) \) at a new and arbitrary test location \( x^* \). We divide \( A \) into \( K \) discrete test locations. Let \( X = \{x_1^*, x_2^*, \ldots, x_K^*\} \) denote the set of test locations where the channel predictions are calculated for the \( M \) APs. Let \( P^m_{RX}(x^*) \) be the mean and \( V^m_{RX}(x^*) \) be the variance of the Gaussian predictive distribution \( p(P^m_{RX}(x^*)|X,y^m,\theta^m,x^*) \) for an arbitrary test location \( x^* \), which can be obtained as

\[
\hat{P}^m_{RX}(x^*) = \mu(x_{AP}^m,x^*) + k^2_{\text{bias}} K^{-1}(y - \mu(X)) \tag{7}
\]
\[
V^m_{RX}(x^*) = k_{\text{bias}}^2 K^{-1} k_{\text{bias}} \tag{8}
\]

where \( \mu(X) = [\mu(x_{AP}^m,x_1),\mu(x_{AP}^m,x_2),\ldots,\mu(x_{AP}^m,x_Q)]^T \) is the mean vector and \( K \) is the covariance matrix of the received powers with entries \( K_{qq'} = C(x_q,x_{q'}) + \sigma^2_{\text{bias}} \delta_{qq'} \), \( k_{\text{bias}} \) is the \( N \times 1 \) vector of cross-covariances \( C(x^*,x_q) \) between the received power at \( x^* \) and at the training locations \( x_q \), and \( k_{\text{bias}} \) is the prior variance, given by \( C(x^*,x^*) \).

Based on the framework described in this section, a SNR map for each AP can be constructed for the quantized space \( X \) of \( A \). Based on the user’s current location and its rate requirement, the proactive approach decides whether the user should stay at the same location or move to a different location to improve its QoS.
B. Centralized Proactive Approach

In the centralized network-centric proactive approach, the network decides which locations are best to the users to improve their QoS. Based on the channel prediction framework, the utility of serving user $n$ by AP $m$ at location $x^* \in \mathcal{X}$ is

$$U_n^m(x^*) = \hat{r}^m(x^*),$$

where $\hat{r}^m(x^*)$ is the predicted rate, associated with predicted receiver power $P_{RX}^m(x^*)$. Then user assignment for the centralized proactive approach can be written as

$$\text{maximize} \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{k=1}^{K} U_n^m(x_k^*) y_{nmk}$$

subject to

$$\sum_{n=1}^{N} \sum_{m=1}^{M} y_{nmk} \leq L_{\max}^m, m \in \mathcal{M},$$

$$d_n y_{nmk} \leq \hat{r}^m(x_k^*), \forall n \in \mathcal{N}, \forall m \in \mathcal{M}, \forall x_k^* \in \mathcal{X}$$

$$\sum_{m,k} y_{nmk} \leq 1, n \in \mathcal{N},$$

$$\sum_{n,m} y_{nmk} \leq 1, 1 \leq k \leq K,$$

$$y_{nmk} \in \{0,1\}, n \in \mathcal{N}, m \in \mathcal{M}, 1 \leq k \leq K.$$ (9f)

Here, the objective (9a) favors a future location and corresponding AP if the associated utility is maximized for the users with its rate requirement; (9b) makes sure the total number of channels to serve by the AP is less than the maximum supported channels of the corresponding AP; (9c) states that each user desired rate requirement should be met; (9d) guarantees that each user is at most assigned a one location and one of the APs; (9e) states that each location cannot be assigned to more than one user; (9f) imposes the binary integer requirements on the optimization variable.

IV. DECENTRALIZED PROACTIVE APPROACH

In the decentralized user-centric approach, the users not only can change AP association, but also move to new locations themselves in discrete time slots (indexed by $t$, where we allow only one user movement$^1$ per time slot) if it improves their utility. In each time slot $t$, the user makes a decision, to stay or not to stay at the current location. If a new location offers a better QoS than the current one, then the user considers moving to that location. However, there is a switch cost associated to it. This switch cost depends on how far the new location is from the current location. The switch cost mimics the user desire to move and is defined as

$$C_s(x_n, x^*) = \alpha_n \|x_n - x^*\|,$$ (10)

where $x_n$ is the current location of the user, $x^*$ is the new location under consideration, and $\alpha_n \geq 0$ is a design parameter that describes the resistance of user $n$ to movement. To ensure that users prefer APs that are less congested, we introduce a congestion penalty. Let $S^m(t)$ be the set of users associated with the AP $m$ at time slot $t$, then the current load of the AP $m$ can be written as $L_{\text{tot}}^m(t) = |S^m(t)|$, where $|$ denotes set cardinality. The congestion penalty for the AP $m$ in time slot $t$ is defined as

$$\lambda^m(t) = \beta^m(L_{\text{tot}}^m(t) - L_{\max}^m),$$ (11)

where $\beta^m$ is a tunable parameter.

The user $n$ is located at $x_n$ during time slot $t-1$ and is served by AP $m'$. The utility function for user $n$ at time slot $t-1$ for the AP $m'$ is defined as

$$U_n^{m'}(t-1, x_n) = \begin{cases} r^{m'}(x_n) - \lambda^{m'}(t-1) & n \in S^{m'}(t-1) \\ 0 & \text{otherwise} \end{cases}$$ (12)

The utility user $n$ would get by moving to location $x^* \in \mathcal{X}$ and associating itself with AP $m$ is given by

$$U_n^m(t, x^*) = \hat{r}^m(x^*) - \lambda^m(t) - C_s(x_n, x^*),$$ (13)

where, when $m = m'$, $\lambda^m(t) = \lambda^m(t-1)$, while for $m \neq m'$,

$$\lambda^m(t) = \lambda^m(t-1) + \beta^m$$

With this formulation, each user runs a local optimization problem to find the best location and AP association that increases its utility and then makes a decision. The corresponding BIP of user $n$ can be written as

$$\text{maximize} \sum_{m=1}^{M} \sum_{k=1}^{K} \left( U_n^m(t, x_k^*) - U_n^{m'}(t-1, x_n) \right) y_{nmk}$$

subject to

$$d_n y_{nmk} \leq \hat{r}^m(x_k^*), m \in \mathcal{M}, x_k^* \in \mathcal{X},$$ (14b)

$$L_{\text{tot}}^m(t-1) - I(m=m') + y_{nmk} \leq L_{\max}^m,$$  \hspace{1cm} $m \in \mathcal{M}, x_k^* \in \mathcal{X},$ (14c)

$$\sum_{m,k} y_{nmk} \leq 1,$$ (14d)

$$y_{nmk} \in \{0,1\}, m \in \mathcal{M}, 1 \leq k \leq K,$$ (14e)

where $I(.)$ is an indicator function which takes value 1 if the argument is true and 0 if it is false. The objective (14a) favors a future location and corresponding AP when the utility gap is maximized; (14b) states that each user’s desired rate requirement should be met; (14c) states that AP $m$ should not be crowded when the user chooses to move to a location that is served by that AP; (14d) guarantees that each user is associated with only one AP and test location; (14e) imposes the binary integer requirements on the optimization variable.

The information that is assumed to be known at the user to solve the optimization problem (14) include: (i) current load at each AP, $L_{\text{tot}}^m$; (ii) maximum supported channels of each AP, $L_{\max}^m$; (iii) SNR map information w.r.t. all APs; (iv) own current utility function.

---

$^1$It is assumed that once the user makes the decision, the updates are done only after the user reaches its new location. Otherwise, there may be discrepancies in terms of the predicted rate, which lead to other user decisions based on inaccurate information. The problem with our sequential approach is that it may take some time to converge and that there should be some sort of coordination among the users.
Fig. 1. User assignment with $N = 10$ and $d_n = 15$ bits/sec/Hz for the reactive approach. The locations of the AP1 and AP2 are shown in blue and magenta diamonds, respectively. The users assigned to AP1 are colored in blue circle and for the AP2 in magenta circle.

Fig. 2. User assignment with $N = 10$ and $d_n = 15$ bits/sec/Hz for the centralized proactive approach. The locations of the AP1 and AP2 are shown in blue and magenta diamonds, respectively. The users assigned to AP1 are colored in blue square and for the AP2 in magenta square.

As mentioned, we only allow one user movement per time slot. The ordering of these movements can be random or in a round-robin fashion.

V. NUMERICAL RESULTS AND DISCUSSION

In this section, we show the performance comparison of the following approaches: (i) the reactive approach, (ii) the centralized proactive approach, (iii) the decentralized proactive approach. The performance metrics are number of satisfied users

$$N_{\text{sat}} = \sum_{m=1}^{M} \sum_{n \in S^m(t)} I(r_m(x_n) > d_n),$$

and the total sum rate of the network

$$R_{\Sigma} = \sum_{m=1}^{M} \sum_{n \in S^m(t)} r_m(x_n).$$

A. Simulation Setup

We consider an indoor region $\mathcal{A}$ of square area $200$ m $\times$ $200$ m consisting of 2 APs. We chose number of users to be $N = 15$ unless otherwise stated. The APs are placed at the locations with coordinates $[0,0]^T$ and $[200,200]^T$. The region $\mathcal{A}$ is discretized in to a grid with 10 m resolution. The set $\mathcal{X}$ contains all the grid locations. We place the users randomly at the grid locations such that each grid contains only one user. A two dimensional true radio propagation field is generated for $\mathcal{X}$ using (1). The predicted radio propagation field is generated for $\mathcal{X}$ based on the $Q$ uniformly distributed training locations using (7). The simulation parameters used to obtain the numerical results are given in Table I.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta$</td>
<td>2.5</td>
</tr>
<tr>
<td>$\sigma_n$</td>
<td>0.01</td>
</tr>
<tr>
<td>$d_c$</td>
<td>5 m</td>
</tr>
<tr>
<td>$L_{\text{max}}$</td>
<td>10</td>
</tr>
<tr>
<td>$M$</td>
<td>2</td>
</tr>
<tr>
<td>$\sigma_\psi$</td>
<td>10 dB</td>
</tr>
<tr>
<td>$K$</td>
<td>400</td>
</tr>
<tr>
<td>$Q$</td>
<td>280</td>
</tr>
</tbody>
</table>

For simplicity, we assume all users have same desired rate requirement $d_n$ and movement resistance $\alpha_n$. Furthermore, we ignore the congestion cost and set $\beta_m = 0, \forall m$. 

TABLE I SIMULATION PARAMETERS
B. User Behavior Example

The user assignment for the three methods is shown in Figs. 1–3. The background rate map is a combined rate map from the two APs. Furthermore, the rate map represents true received power for the reactive approach and predicted received power for other two methods. For the reactive approach, the network performs user to AP associations such that the user rate requirement is met. For example in Fig. 1 with \( d_n = 15 \) bits/sec/Hz, user 4 and 5 are not assigned to any APs, as their rate requirement is not met due to their locations in the poor coverage areas. For the proactive centralized approach, the network assigns locations with best rates to users such that each user’s rate is maximized. It is easy to observe in Fig. 2 that now users have been moved to locations with highest possible rate and all user rate requirements are met. In Fig. 3, the initial reactive approach assignment and the final decentralized proactive assignment is shown. Based on the switch penalty factor, the users move to locations that improve their rate. In the initial reactive assignment, user 4 and 5 do not meet their rate requirement. The user in the decentralized approach moves to a location that meets its rate requirement. In this case, all users’ rate requirements are satisfied after moving to new locations.

C. Quantitative Analysis

Impact of desired rate \( d_n \): The impact of the desired rate on \( N_{sat} \) and \( R_S \) for the three approaches with true and predicted fields is shown in Figs. 4–5. When \( d_n \) is increased, the number of satisfied users decreases and ultimately reaches zero (see Fig. 4). As a result, \( R_S \) decreases with increase in \( d_n \) and finally becomes zero (see Fig. 5). This is expected as with increase in \( d_n \), it is always impossible to satisfy all users demand requirements. The reactive approach is most sensitive to increase in \( d_n \): \( N_{sat} \) starts decreasing after \( d_n = 13 \) bits/sec/Hz. It should be noted that the \( N_{sat} \) and \( R_S \) curves for the reactive approach with true and predicted fields are on top of each other. For the case of making assignments based on the predicted field, the centralized proactive approach is more robust to increase in \( d_n \) when compared to the decentralized approach. However, the trend is the same for both approaches when they have access to the true field. It can be observed easily that user assignment based on true field offers more robustness with increase in \( d_n \). Furthermore, it is evident that user assignment based on true field offers better \( R_S \).

Impact of switch penalty factor \( \alpha_n \): In Figs. 6–7, we depict the impact of the switch penalty factor \( \alpha_n \) on \( N_{sat} \) and \( R_S \) for the three approaches with true and predicted fields. An increase in \( \alpha_n \) means the switching cost associated

---

2At each location, maximum rate is chosen from the rates provided by each AP.
with movement is higher. As expected, for the decentralized proactive approach, an increase in $\alpha_n$ results in a decrease in $N_{\text{out}}$ and $R_S$. This is due to the fact that users are more reluctant to move to new locations to increase their rate. In this case also, for decentralized proactive approach the $R_S$ of predictive and true fields coincide for large $\alpha_n$. It should be noted that there is no impact of $\alpha_n$ for the reactive and centralized proactive approaches. Finally in our simulation results, we observe that users do not keep moving to better locations, instead they converge in a few iterations to locations that meet their desired rate requirement.

VI. CONCLUSION

In this paper, we proposed a location-aware user-centric proactive resource allocation approach, called LAPRA. The proactive approach makes use of channel predictions by using location information. We analyzed and compared our approach with a classical reactive approach. The reactive approach suffers from poor QoS, while LAPRA improves the QoS by offering higher user satisfaction and higher total sum rate offered by the network. We have considered the availability of perfect location information during channel prediction. The analysis of uncertain location information is left as future work.

ACKNOWLEDGMENT

This research was supported in part, by Bell Lab internship, and by the European Research Council, under Grant No. 258418 (COOPNET). Discussions with Dr. Themistoklis Charalambous are gratefully acknowledged.

REFERENCES


