J. Indian Inst. Sci.
A Multidisciplinary Reviews Journal
ISSN: 0970-4140 Coden-JIISAD

REVIEW ARTICLE

Mobility-Driven Association Policies for Dense Wireless Networks

Pranav Madadi, François Baccelli and Gustavo de Veciana*

Abstract | The primary approach to increase coverage and capacity in infrastructure-based wireless networks is densification. Densification, however, presents a major challenge when serving mobile users, which is the overhead associated with the increased rate of base station handovers. Assuming a prior knowledge of a mobile's trajectory and base stations' locations, we formulate the problem of determining the sequence of handovers that optimize the trade-off between the mobile user's perceived throughput and handover overheads in noise-limited environments. Under appropriate conditions, we show that the problem reduces to determining a maximum weight path in a directed acyclic graph induced by the mobile user's trajectory. In practice, knowledge of a mobiles' trajectory may be limited and one may also want to limit the handover complexity, whence we propose a new class of mobility-driven greedy association policies. The greedy policies are based on defining a handover support set, which constrains both the possible handovers and the complexity/information requirements. In a setting where base station locations follow a Poisson point process, we show that the performance of such handover processes follows a continuous-time Markov process which can be analyzed using complex variable techniques. This enables one to explore the optimization size/shape of the handover support set for mobility-driven greedy handover strategies and their relative performance compared to traditional association policies.

1 Introduction

The high rate of technological advancement in the wireless industry has resulted in a substantial increase in the use of mobile devices such as smartphones and tablets. Indeed, cellular traffic generated by mobile devices has grown 18-fold in the past 5 years and by 63% in 2016 alone[1] and is expected to increase 1000-fold in the next decade. One can expect intensive use of mobile devices (smartphones) during commute time. Applications such as video/audio streaming accounted for 60% of total mobile data traffic in 2016[1]. Geolocation services/maps and social media are also extensively used during commute periods. In the future, with an increase in the use of public transportation and the emergence of self-driving cars, this volume could grow substantially.

Mobile users expect excellent coverage with high data rates for their applications to run smoothly anytime/anywhere. To meet the increasingly stringent performance requirements of users along with the increase in traffic volume, various technologies are being explored for 5G networks. Along with techniques like millimeter wave and massive MIMO, future networks are likely to be ultra-dense with deployment of a large number of small cells to boost capacity and/or achieve seamless coverage. Thus, a further challenge for future 5G wireless networks will be delivering high capacity to large numbers of

[1] Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, USA. *pranav93@utmail.utexas.edu

highly mobile users and the association/handover problem will be of great interest.

The high frequency of handovers resulting from mobility in dense networks can be detrimental to user performance. Unfortunately, the negative impact of densification while serving mobile users is often ignored[2]. A high handover rate results in high overhead for both applications and infrastructure. In addition to the handover signaling overheads, the handover procedure interrupts the data flow to the user due to link termination with the serving base station and link establishment with the target base station. A recent study[3] showed that handovers generally cause short-term disruptions in various applications. Similarly, simulation studies[4] have shown that handovers degrade the performance of real-time applications such as VoIP. Thus, the increased rate of handovers diminishes the benefits of base station densification.

Modeling and improving handover performance have been extensively addressed in the cellular network literature. The emphasis was on offloading the mobile user traffic to macrocells in heterogeneous networks to reduce the frequency of handovers, but the increase in the volume of mobile users then leads to heavy congestion at macro-cells.

A handover management technique based on self-organizing maps is proposed in[5] to reduce unnecessary handovers for indoor users in two-tier cellular networks. The authors in[6] present a study to avoid unnecessary vertical handovers and reduce the overall packet delay for low-speed users in two-tier downlink cellular networks. Handover signaling overhead reduction algorithms are proposed in[7] for two-tier networks and in[8] for cloud-RAN-based heterogeneous networks. Handover delay for a single-tier network is characterized in[9]. However, none of the aforementioned studies tackle the interplay between handover cost and capacity gain as a function of the base station intensity.

In this paper, we explore a class of mobility-driven association policies that optimize the trade-off between handover cost and throughput gains. There have been many studies of user association policies other than the closest BS policy for different spatial models such as single-tier (homogeneous) and multi-tier (heterogeneous) networks, see[10–16] and references therein. Primarily, they can be classified into three categories based on the metrics considered for association: (1) received signal strength, i.e., highest SINR (or highest SIR) in interference-limited networks, (2) load balancing (for heterogeneous networks, it is often coupled with inter-cell interference mitigation[17]) and, (3) energy aware, where one considers the operational power consumption of base stations as the user association metric so that users are likely to be associated with energy-efficient base stations[18].

Most of these works accurately model spatial traffic variations but none consider user mobility. In the context of multi-hop ad hoc networks, a selection of a receiver for relaying the information that is aware of the direction of the destination is studied in[19].

We model the cellular network using stochastic geometry, which is a widely accepted mathematical tool to model and analyze cellular networks which enables performance characterization in terms of the base station intensity, λ as well as other physical layer parameters (see[20] for a survey). In such networks, under the closest base station association policy, the intensity of handovers scales very poorly, i.e., proportionally to $\sqrt{\lambda}$[21]. Thus, our proposed association policies are also geared at reducing the scaling of the intensity of handovers.

The handover rate for Poisson cellular networks in terms of the base station intensity is characterized in[21] for a single-tier scenario and in[22] for multi-tier scenarios. However, none of the aforementioned studies investigate the integrated effect of network densification (i.e., BS intensity) in terms of both the handover cost and the throughput gains. The authors in[23, 24] proposed a simple handover skipping scheme to mitigate handover cost in single-tier networks that improves user throughput at higher velocities. They advocate sacrificing the best base station association to skip some handovers along the user trajectory. This proposed skipping scheme is topology agnostic and can result in non-efficient skipping decisions. In[23], the authors exploited topology awareness and user trajectory estimation to propose smart handover management schemes in single and two-tier downlink cellular networks. However, it is difficult to conduct tractable analysis for the proposed handover skipping schemes due to the random shape of the Voronoi cell and the random location and orientation of the trajectory within the Voronoi cell.

Contributions In this paper, we study the problem of determining optimal base station association policies that trade-off throughput with handover penalties. We consider a noise-limited environment where the base stations are distributed according to a Poisson point process.

We model the cost of a handover by considering the loss in throughput associated with handover delays. We also introduce an additional fixed loss in the data per handover to model the packet loss and signaling overhead.

In our first attempt, we consider a setting where the mobile user's trajectory is known as well as the direction of motion and the location of all base stations in space. In this setting, we consider the total volume of data delivered to the mobile user accounting for handover overheads during the mobile's trajectory as our performance metric. We show that the problem of optimizing this metric is challenging and, under some simplifying assumptions, develop a dynamic programming problem that reduces the above optimization problem to find the maximum weight path in a directed acyclic graph.

Given users often have limited information about the base station, we consider the setting with high handover costs and constrained information about the base station locations by defining a handover support set. In this case, we propose a greedy approach where the user associates with the base station that is farthest in the direction of its motion, i.e., greedily pick the base station that results in the maximum connection time. The constraint on the knowledge of base stations also limits the greediness. We model the handover support set of the tagged user by a geometric region, i.e., all the base stations within a specific set centered at the user. Here, we consider the time average of the throughput as the performance metric and optimize it with respect to parameters associated with the geometric region.

Under our proposed greedy association policy, using the properties of the Poisson point process, we show that the evolution of the mobile user's association base station is Markovian over time. More precisely, we establish a connection between the random geometry of base stations and the theory of Markov processes with respect to time. We then study various properties of this continuous state space Markov process such as its irreducibility, aperiodicity, and characterize its stationary distribution for one specific class of association policies.

Using various results on Markov processes, we evaluate the average throughput seen by the mobile user in closed form. We show that there exists an optimal handover support set, i.e., optimal size of the given geometric region maximizing the average throughput and evaluate it. We then compare the performance of our greedy strategy with the optimal performance of the dynamic programming solution and the closest base station association policy.

We then implement the policy that is close to the current technology complying with the 3GPP specifications. We consider a mobile user moving according to a random way point model and periodically evaluate its SNR value from the serving and neighboring base stations in the presence of fading. We then evaluate its average rate incorporating the penalty of handovers and compare this performance with our proposed model.

The class of models studied in the present paper has several structural parameters and an exhaustive study of all cases is not possible in a single paper. Some of these parameters come from physics like the attenuation function. Others from the type of wireless network considered, as the expression for the Shannon rate, which depends on whether the network is noise or interference limited. The other parameters are simple parameters of the algorithms, e.g., the shape or the dimension of the handover support set. In all cases, we decided to mainly focus on the structural parameters leading to the simplest analytical expressions. For instance, we focus on noise-limited networks and on rectangular handover support set. We also show that the techniques apply to more general cases (e.g., more general shapes) but leave a detailed analysis of these other cases for future research. In particular, the interference-limited case will require significant further work.

Paper Organization The paper is organized as follows. We describe our system model in Sect. 2 and define the set of accessible base stations. In Sect. 3 we first formulate the problem of finding the optimal sequence of base station associations and handover times and then pose constraints that reduce the problem to one that can be solved via dynamic programming. In Sect. 4, we propose our mobility-driven greedy association policies and study the problem as a parametric optimization of the underlying handover geometry. In Sect. 5, we compare the performance of association policies in various scenarios. Section 6 concludes and describes future work.

2 System Model

Consider a dense noise-limited wireless network with base stations denoted through their locations on the Euclidean plane. The configuration

Table 1: Table of notation for the dynamic programming

Symbol	Definition
$\phi = \{b_1, b_2, \dots\}$	Set of base stations
T	Mobile user trajectory of motion
$\mathcal{B}(b_i)$	Handover set of base station b_i
$\left[u_{b_i}^{(1)}, u_{b_i}^{(2)}\right]$	Interval during which the user can associate with base station b_i
$\mathbf{p} = (p_1, p_2, \dots, p_k)$	Association sequence of length k
t_i	Handover time from p_i to p_{i+1}
$\mathbf{t} = (t_1, t_2, \dots, t_{k-1})$,	Handover times for association path $\mathbf{p}$
$r(p_i, p_{i+1})$	Net rate associated with the handover $p_i \to p_{i+1}$
δ, c	Time of service interruption due to a handover and a fixed cost

of the base stations, $\Phi = \{B_1, B_2, \dots\}$, is assumed to be a realization of a Poisson point process of intensity λ in $\mathbb{R}^2$. In the sequel, we let B_i denote both the location of the base station and a state representing the user's association with base station B_i. We consider downlink transmissions and assume that all base stations transmit at a fixed power ξ.

We adopt a deterministic isotropic bounded path loss model represented by a function $l(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$. Specifically, we consider a special case of the path loss functions, i.e., the stretched exponential, which are proven to be a good model for short-range propagation in many environments[25],

$$l(x) = \exp(-\kappa x^{\beta}), \tag{1}$$

where κ, β are fitting parameters. Later, in the paper we consider various values of parameter β to model the different attenuation in path loss functions and to derive closed-form solutions.

We consider a tagged mobile user moving at a constant velocity v along a trajectory, T, starting from the origin o and moving to a destination d. Initially we assume that the user is aware of the base station locations, denoted by $\phi = \{b_1, b_2, ...\}$. We assume that the narrow band beamforming mitigates the interference and consider SNR while evaluating the Shannon rate of the mobile user. We further, incorporate the impact of handovers on the rate seen by the mobile user.

In general, handovers are performed in three phases: initiation, preparation, and execution[26]. This involves signaling overhead between the user, the associated base station, the target base station and the core network, which often interrupts the data flow to the user. We denote the duration of each interruption (handover delay) by δ seconds, measured from the beginning of the initiation phase to the end of the execution phase. Such delays can be significant[27]. In the sequel, we model the cost of a handover by considering the loss in throughput associated with such handover delays. Further, we also introduce an additional fixed loss of c bits per handover to model the packet loss and signaling overhead (Table 1).

In the sequel, we adopt the notation $||\cdot||$ to denote Euclidean distance and $[\cdot]^+$ which implies that

$$[a]^+ = \begin{cases} a & \text{if } a > 0, \\ 0 & \text{if } a \leq 0. \end{cases} \tag{2}$$

3 The Dynamic Programming Approach

3.1 Mobility-Driven Association Policies

In this subsection, we introduce the general notion of mobility-driven association policies that represent feasible associations encountered by a typical tagged mobile user.

The tagged user moving along its trajectory goes through different association states, p_i which corresponds to either being connected (associated base station) or disconnected as illustrated in Fig. 1. Let $\mathbf{p} = (p_1, p_2, \dots, p_k)$ denote the sequence of such association states of certain length $k \in \mathbb{N}$ and $\mathcal{P}$ denote set of feasible association sequences.

The feasibility of an association sequence depends on the tagged mobile user's:

- *Trajectory* We assume that the user is moving along a straight line starting from origin o and moving to a destination d.

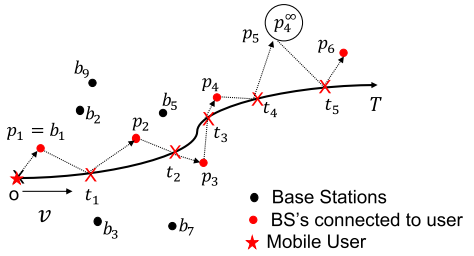

Figure 1: Different associations states a tagged mobile user has while moving along its trajectory T at a constant velocity v.

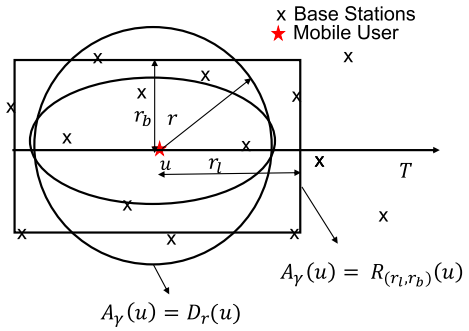

Figure 2: Different possible handover support sets for a mobile user at u, $A_\gamma(u)$.

Without loss of generality, we assume $T = \{u(t) = vt : t \in [0, t(d)]\}$, where $u(t)$ denotes the user's location on the trajectory at time t and $t(u)$ denotes the time at which the tagged user reaches the location u on the trajectory.

- *Possible associations* We assume that at any given time the mobile user can only connect to a base station from a set of accessible base stations defined below. The mobile user is assumed to be disconnected in the case where there are no accessible base stations. Further, we assume that if the mobile user can be connected, it will be connected, i.e., the user is disconnected only in the absence of accessible base stations. This is in part inspired by the threshold service model considered in[28], where the user is idle in the case where the SNR received is less than a given threshold.

Accessible base stations For the mobile user at location u on T, we define the set of its accessible base stations, $\Phi_\gamma(u)$, with parameter γ, as

$$\Phi_\gamma(u) = \Phi \cap A_\gamma(u), \tag{3}$$

where $A_\gamma(u)$ is the handover support set defined here as a convex region centered at u and with parameter γ. The parameter γ can either be a scalar (e.g., when $A_\gamma(u)$ is a disc $D_r(u)$ of radius r, $\gamma = r$) or a vector (e.g. when $A_\gamma(u)$ is a rectangle $R_{(r_l,r_b)}(u)$ centered at u, $\gamma = (2r_l, 2r_b)$ represents the side lengths of the rectangle; for the case of an ellipse as illustrated in Fig. 2, γ will be the parameters of this ellipse).

Disconnected states We shall distinguish various disconnected states for purposes that will be clear in the next subsection. If a user is associated with a base station, $b_i \in \phi$ and becomes disconnected, we denote it as being in state b_i^∞. Similarly, if the user is disconnected at the beginning of motion, i.e., at the origin o, we denote it as being in state b_o^∞. Let $\psi = \{b_o^\infty\} \cup \{b_i^\infty | b_i \in \phi\}$ denote the set of possible disconnected states.

Maximal connection intervals We define the longest connection interval for a base station, b_i, as the longest line segment on the user's trajectory on which the tagged user can associate with base station b_i. This interval is denoted by $\left[u_{b_i}^{(1)}, u_{b_i}^{(2)}\right]$, where $u_{b_i}^{(1)}, u_{b_i}^{(2)}$ are locations on the user's trajectory, T. Figure 3 illustrates the interval for a handover support set equal to $D_r(u)$. In this case $u_{b_i}^{(1)}, u_{b_i}^{(2)}$ are the intersections of the circle $D_r(b_i)$ with the user's trajectory T.

Handover sets Let us define the handover set, $\mathcal{B}(b_i)$ as the set of base stations to which the mobile user could realize a handover from base station $b_i \in \phi$, i.e., all the base stations that are accessible from some point on the maximal connection interval of b_i, $\left[u_{b_i}^{(1)}, u_{b_i}^{(2)}\right]$. Further, for any disconnected state b_i^∞ the handover set $\mathcal{B}(b_i^\infty)$ is the first available base station along the trajectory:

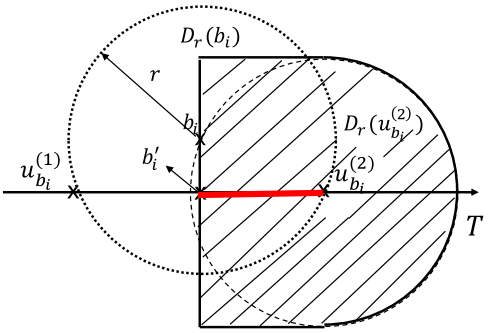

Figure 3: Figure illustrating the handover sets for a handover support set of a disc of radius r denoted as $D_r(u)$.

$$\mathcal{B}(b_i^{\infty}) = \left\{ b_j \in \phi \ \text{ s.t. } \ u_{b_j}^{(1)} = \min_{b_k \in \phi} \left[u_{b_k}^{(1)} : u_{b_k}^{(1)} > u_{b_i}^{(2)} \right] \right\}. \tag{4}$$

Given the properties of the Poisson point process, this is a singleton almost surely. Note that the mobile user enters the disconnected state from a base station b_i only when $\mathcal{B}(b_i)$ is empty.

Recall that $\phi_r(o), \phi_r(d)$ denote the set of accessible base stations for the tagged user at the origin o and the destination d, respectively. Then, for a given trajectory and possible associations for the tagged user, we now formally define feasible association sequences.

Definition 1 A feasible association sequence, $\mathbf{p} = (p_1, p_2, \ldots, p_k)$, of length $k \in \mathbb{N}$, for the trajectory T is such that:

- p_1 is either a base station in $\phi_r(o)$ or the disconnected state b_o^{∞}, when $\phi_r(o)$ is empty,
- for each $i > 1$, p_i is either a base station in $\mathcal{B}(p_{i-1})$ or the disconnected state p_{i-1}^{∞}, when $\mathcal{B}(p_{i-1})$ is empty, and
- $p_k \in \phi_r(d)$ or p_{k-1}^{∞} if $\phi_r(d)$ is empty.

Given an association sequence, $\mathbf{p} \in \mathcal{P}$ of length k, one can choose a sequence of handover times $\mathbf{t} = (t_1, t_2, \ldots, t_{k-1})$, where t_i denotes the handover time from p_i to p_{i+1}. Given $\mathbf{p}$, there are constraints on $\mathbf{t}$ for sake of simplicity, that we will describe later in Assumption 2. Let $\mathcal{T}_{\mathbf{p}}$ denote the set of all such feasible sequences of handover times for $\mathbf{p}$. Thus, a specific mobility-driven association policy is equivalent to selecting a pair $(\mathbf{p}, \mathbf{t})$, with $\mathbf{p} \in \mathcal{P}$ and $\mathbf{t} \in \mathcal{T}_{\mathbf{p}}$.

For each pair $(\mathbf{p}, \mathbf{t})$ with $\mathbf{p} \in \mathcal{P}$ and $\mathbf{t} \in \mathcal{T}_{\mathbf{p}}$, let $r(\mathbf{p}, \mathbf{t})$ be the total volume of data delivered to the tagged mobile user during its trajectory T. The amount of data received in each association state can be estimated by considering the Shannon rate of the user when in this state and by taking into account the interruptions in data transfers due to handovers. In the sequel, we evaluate $r(\mathbf{p}, \mathbf{t})$ as the sum of the amount of data received in each association state.

3.2 Restrictions Leading to Dynamic Programming

The general focus of the present paper is about association policies $(\mathbf{p}, \mathbf{t})$ optimizing the reward function $r(\mathbf{p}, \mathbf{t})$ in this context. This general optimization problem is quite complicated, primarily due to the continuum of possible handover times, and it is not solved in the present paper. We nevertheless describe it in more detail in Appendix A for the sake of completeness. In this subsection, we propose two natural restrictions on $(\mathbf{p}, \mathbf{t})$ that allow one to reduce this optimization problem to dynamic programming.

Definition 2 The forward handover set $\mathcal{B}^f(b_i)$ of b_i is the set of base stations b_j in $\mathcal{B}(b_i)$ such that $\Pi_T(b_j) > \Pi_T(b_i)$, where $\Pi_T(b_i)$ denotes the projection of a base station location, $b_i \in \phi$ on the trajectory T.

Given the set of base stations is a realization of a homogeneous Poisson point process, all projections of the base station locations on T are almost surely unique, i.e., no two projections are the same. In words, handovers from b_i can only be to a base station whose projection on T is encountered after the projection of b_i on T.

Assumption 1 We restrict the set of feasible association sequences to forward association sequences. Such sequences are obtained when replacing the handover sets $\mathcal{B}(\cdot)$ in Definition 1 by forward handover sets $\mathcal{B}^f(\cdot)$.

The base stations in $\mathcal{B}^f(b_i)$ are in the set illustrated in Fig. 3. Note that the definition for the handover set corresponding to the disconnected state b_i^{∞} is the same as before (4) and that a forward association sequence, $\mathbf{p} = (p_1, p_2, \ldots, p_k)$ is an association sequence such that, for every $i > 1$, p_i is either a base station in $\mathcal{B}^f(p_{i-1})$ or the disconnected state p_{i-1}^{∞} if the latter is empty. Let $\mathcal{P}_f \subset \mathcal{P}$ denote the set of all possible forward association sequences.

Definition 3 For a given trajectory, T, set of base stations, ϕ, and associated disconnected states, ψ, we define the forward association graph $G = (V, E)$ as follows: $V = \phi \cup \psi \cup \{o, d\}$ and (p_i, p_{i+1}) belongs to E, if (p_i, p_{i+1}) belongs to some forward association sequence in $\mathcal{P}_f$. Similarly, the edge (o, p_1) (resp. (p_k, d)) belongs to E if p_1 (resp. p_k) is the first (resp. last) state in some forward association sequence in $\mathbf{p} \in \mathcal{P}_f$.

The main justification for Assumption 1 is that the forward association graph G is a directed acyclic graph (DAG). Further, the forward association sequences are the paths from node o to node d in this graph. Figure 4 gives an instance of such a graph.

As already explained, our aim is to evaluate the reward, $r(\mathbf{p}, \mathbf{t})$ associated with the forward

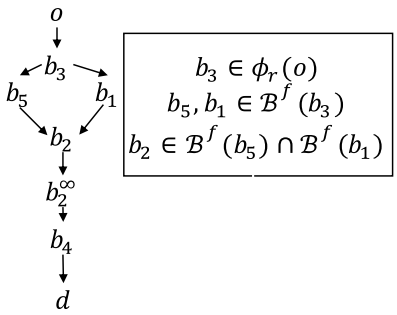

Figure 4: An example of directed acyclic graph.

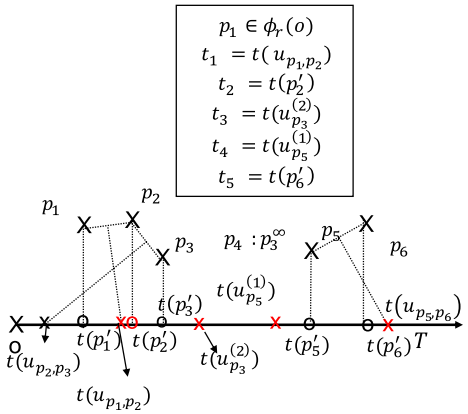

Figure 5: Figure illustrating all cases of handover times for a mobile user moving along a straight line at unit velocity.

association path $\mathbf{p}$ by representing it as the sum of rewards on the edges along the path. We do this by further restricting ourselves to the case where the sequence of handover times, $\mathbf{t}$, is a function of the association sequence $\mathbf{p}$. We select natural handover times from the continuum under the constraint that $t_i \in \mathbf{t}$ depends only on the pair $(p_i, p_{i+1}) \in \mathbf{p}$.

Recall that $\left[u_{b_i}^{(1)}, u_{b_i}^{(2)}\right]$ is the maximal connection interval for base station b_i. Also recall that if the tagged user can be connected, it will be connected and that a handover from p_i to the disconnected state happens at location $u_{p_i}^{(2)}$. Similarly, a handover from the disconnected state to base station p_{i+1} happens at location $u_{p_{i+1}}^{(1)}$.

The main justification for Assumption 2 below is illustrated in Fig. 5 and is the following: for a handover from base station p_i to p_{i+1}, it is best, in terms of Shannon rate, to handover at a time where the mobile is equidistant to both base stations. Let $u_{p_i,p_{i+1}}$ denote the location on the trajectory T when the user is equidistant from both the base stations, i.e., $||p_i - u_{p_i,p_{i+1}}|| = ||u_{p_i,p_{i+1}} - p_{i+1}||$. Let p_i' be a short notation for $\Pi_T(p_i)$. If the time $t(u_{p_i,p_{i+1}})$ is within the range $\left[t(p_i'), t(p_{i+1}')\right]$, then the best time for the handover $p_i \to p_{i+1}$ is $t_i = t(u_{p_i,p_{i+1}})$. If $t(u_{p_i,p_{i+1}})$ is not in this range, under the constraint that it should depend only on p_i, p_{i+1}, it is best to handover at one of the extreme points of the interval $\left[t(p_i'), t(p_{i+1}')\right]$. More precisely, we make the following additional assumption:

Assumption 2 We assume that the handover time from p_i to $p_{i+1} \in \phi$ is given by:

$$t_i = \begin{cases} t(p_i') & \text{if } t(u_{p_i,p_{i+1}}) < t(p_i'), \\ t(p_{i+1}') & \text{if } t(u_{p_i,p_{i+1}}) > t(p_{i+1}'), \\ t(u(p_i,p_{i+1})) & \text{if } t(u_{p_i,p_{i+1}}) \in [t(p_i'), t(p_{i+1}')], \end{cases} \tag{5}$$

and that in addition

$$t_i = \begin{cases} t(u_{p_i}^{(2)}) & \text{for } p_i = b_j \to b_j^{\infty}, \\ t(u_{p_{i+1}}^{(1)}) & \text{for } b_j^{\infty} \to p_{i+1}. \end{cases} \tag{6}$$

3.3 Rewards

Having defined handover times as a function of the forward association sequence, the association reward, $r(\mathbf{p}, \mathbf{t})$ now depends only on the association sequence and is denoted by $r(\mathbf{p})$. Let $s_{p_i}(t)$ be the Shannon rate seen by the user at any given time t

$$s_{p_i}(t) = \omega \ln\left(1 + \frac{\xi \times l(||p_i - u(t)||)}{\sigma^2}\right), \text{ for } t_{i-1} \le t < t_i, \tag{7}$$

where σ^2 is the noise power, ω is the system bandwidth and $||\cdot||$ denotes Euclidean distance. Since the rate seen by the user in the disconnected state is zero, we set $s_{p_i}(t) = 0$, if $p_i \in \psi$.

Net rate associated with a handover We define the net rate associated with a handover $p_i \to p_{i+1}$, $r(p_i, p_{i+1})$ as the total volume of data received by the mobile user while moving from $p_i' \to p_{i+1}'$. Then under Assumptions 1 and 2, the association reward, $r(\mathbf{p})$ for a forward association sequence of length k, is additive, and with $p_0 = o, p_{k+1} = d$, we have $r(\mathbf{p}) = \sum_{i=0}^{k} r(p_i, p_{i+1})$, where,

Table 2: Table of notation for greedy association

Symbol	Definition
$\Phi_\gamma(u)$	Accessible base station for mobile user at u
$(P(t), t > 0)$	Random process denoting the association base station to the mobile
$(D(t), t > 0)$	Random process denoting the distance to the association base station
$(S(t), t > 0)$	Random process denoting the Shannon rate seen by the mobile user
$(T_i)_{i=1}^{\infty}$, $u(T_i)_{i=1}^{\infty}$	Handover times and the location of the mobile user
$\mathbf{Z}_n = (X_n, Y_n)$	Relative position of the association base station w.r.t. mobile user's location $u(T_n)$
ρ	Rate of handovers
$\mathcal{S}(\gamma)$	Effective throughput

$$r(p_i, p_{i+1}) = \left[\left(\int_{t(p_i')}^{t_i-\delta/2} s_{p_i}(t)\mathrm{d}t + \int_{t_i+\delta/2}^{t(p_{i+1}')} s_{p_{i+1}}(t)\mathrm{d}t\right) - c\right]^+, \tag{8}$$

and where the handover times are defined as prescribed in Assumption 2.

As already mentioned, we model the cost of a handover by assuming that data transmission is interrupted for δ seconds, equally split between the two states involved in handover. Further, the additional fixed loss of c bits in the data per handover is modeled by subtracting c bits from the total volume of data delivered from each base station. This captures packet losses and signal overhead. In the following two specific cases,

- the value of the upper limit is less than the lower limit, $t(p_{i+1}') - \delta/2 < t_i$, i.e., when the time taken for handover (δ) is more than the time the user is associated with a base station,
- the volume of data delivered is less than the penalty c, so the total volume of data delivered is assumed to be equal to zero.

3.4 Optimal Forward Association Policy

Under Assumptions 1 and 2, the problem of finding the maximum total data delivered to the user reduces to:

Problem 1 *Determine the optimal reward*

$$r_f^* = \max_{\mathbf{p}}\{r(\mathbf{p}) : \mathbf{p} \in \mathcal{P}_f\}, \tag{9}$$

where $\mathcal{P}_f$ denotes the set of all forward association paths from o to d.

Thanks to the observations of the previous subsections, this is equivalent to finding the largest weight path from node o to node d in the DAG of Definition 3, which can be solved using a classical dynamic programming approach[29].

Given that ϕ is a realization of a Poisson point process of intensity λ, the mean number of base stations seen by the tagged mobile user along its trajectory and that are within distance r from its path is equal to $2rd\lambda + \pi r^2\lambda$, where d is the total distance traversed by the user. The number of nodes in the directed acyclic graph is equivalent to number of base stations, thus $|V| = O(2rd\lambda + \pi r^2\lambda)$. Now, a given base station can have edges to base stations in the region as illustrated in Fig. 3, thus we have $|E| = O(\lambda^2 r^3 d)$. Thus, the mean time complexity of solving the optimal path in the graph is $O(|V| + |E|) = O(\lambda^2 r^3 d)$[30].

Thus, Problem 1 can be solved (using dynamic programming) with a linear time complexity with respect to the total distance traversed. Considering a certain realization of base stations, we solve the problem using Dijkstra's algorithm[30]. The performance analysis of this approach is further studied in Sect. 5 (Table 2).

4 Farthest Greedy Policy

4.1 Setting and Definition

This section is focused on policies that have only partial (local) knowledge on the locations of base stations. We consider networks with high handover costs and propose association policies where the mobile user is constrained to associate with a base station in its handover support set and stay connected to this base station for the maximum amount of time, i.e., till it is no longer in its handover support set. Under this constraint, we propose a policy where the user greedily associates with the base station that maximizes the connection time, i.e., the base station that is farthest out in the direction of its motion. Thus, the handover support set around the mobile represents both a

constraint on the information available as well as the degree of greediness in avoiding handovers. We analyze these policies using Markov chains. Since associating with a base station that is at a large distance may result in poor throughput, we then optimize the performance of such greedy association policies with respect to the size of the handover support set in the second step.

Note that unlike in the previous section the mobile user is not aware of the locations of all the base stations along the trajectory of its motion. Thus, we represent the set of base stations as Φ and still assume that the base stations, $\{B_i\} \in \Phi$, are a realization of a Poisson point process of intensity λ and that the tagged user is moving along the x-axis, starting from the origin, i.e., $T = \{u(t), t \geq 0\}$. As already mentioned, to evaluate the performance of greedy policies, we establish a connection with the theory of Markov processes. The steady state of these processes are analyzed. Thus, we assume that the tagged user's trajectory has infinite length.

We model the constraints on possible handovers and the information by assuming that, at any location u on T, the tagged mobile user is only aware of the set of accessible base stations $\Phi_\gamma(u)$ that are within the handover support set $A_\gamma(u)$ as defined in (3).

In this section, we use a single disconnected state denoted by b^∞. Then, at $t = 0$, the mobile user is either in the disconnected state b^∞ or connected to a base station in $\Phi_\gamma(o)$. An association process, $(P(t), t \geq 0)$, is a piece-wise constant, right-continuous stochastic process with state space $\Phi \cup b^\infty$ denoting the association state of the tagged mobile user at time t. This process is assumed to have a countable set of discontinuities $(T_n)_{n=1}^\infty$ (T_n is the handover time from station $P(T_n)$ to $P(T_{n+1})$), without accumulation points, and to be such that

$$P(t) \in \begin{cases} \{\Phi_\gamma(o), b^\infty\}, & \text{for } 0 \leq t < T_1, \\ \{\mathcal{B}(P(T_n)), b^\infty\}, & \text{for } T_n \leq t < T_{n+1}. \end{cases} \tag{10}$$

Below, we retain the definitions of maximal connection intervals, $\left[u_{B_i}^{(1)}, u_{B_i}^{(2)}\right]$ and handover sets $\mathcal{B}(B_i)$ for a base station $B_i \in \Phi$ from the previous sections. Recall also that $\Pi_T(B_i)$ denotes the projection of the base station's location on trajectory T. We define the farthest greedy policy as follows:

Definition 4 The farthest greedy policy is the policy where,

1) when a handover happens, the mobile user chooses as next station the accessible base station that is farthest out in the direction of its motion, i.e., for $T_0 = 0$ and $n \geq 1$,

$$P(T_n) = \underset{B_i \in \Phi_\gamma(u(T_n))}{\arg\max} \times [|\Pi_T(B_i) - u(T_n)| : \Pi_T(B_i) > \Pi_T(P(T_n))], \tag{11}$$

2) the mobile stays connected to this base station for the maximum amount of time, i.e.,

$$T_{n+1} = t\left(u_{P(T_n)}^{(2)}\right). \tag{12}$$

4.2 Reward as Effective Throughput

For all association process $(P(t), t \geq 0)$, one can associate a distance process $(D(t), t \geq 0)$ denoting the distance from the mobile user to its association base station when connected and infinity otherwise given by:

$$D(t) = \begin{cases} \infty \text{ for } P(t) = b^\infty, \\ ||u(t) - P(t)|| \text{ otherwise}. \end{cases} \tag{13}$$

Figure 6 illustrates a realization of association and distance processes.

The associated SNR of the Shannon rate process, $(S(t), t \geq 0)$ is then given by:

$$S(t) = \omega \ln\left(1 + \frac{\xi \times l(D(t))}{\sigma^2}\right). \tag{14}$$

By analogy with what was done in the last section, we now have the following natural definition:

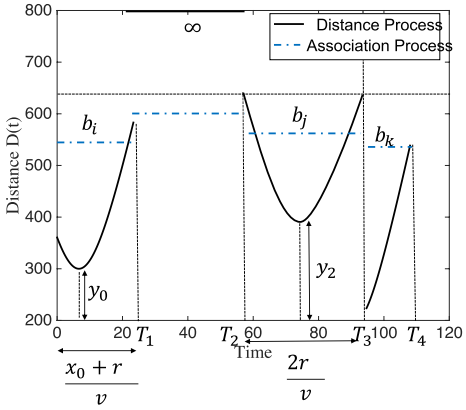

Figure 6: Distance from the association base station to the mobile user.

Definition 5 The effective throughput $\mathcal{S}(\gamma)$ seen by a mobile user is defined as the long time average, when it exists, of the Shannon rate seen by the user when its service is interrupted for δ seconds and penalized c bits for every handover.

4.3 Markov Formulation

For simplicity, in the sequel, we consider the case where the handover support set is a fixed rectangular shape, $A_\gamma(u) = R_{(r_l,r_b)}(u)$ centered at the user location u. Note however that the analysis can be extended to other shapes like ellipses.

Definition 6 At any given handover time, T_n, we define the relative position as the coordinates, $\mathbf{Z}_n = (X_n, Y_n)$ of the association base station, $P(T_n)$ with respect to the location $u(T_n)$ on the tagged users' trajectory T, if the user is connected and infinity otherwise:

$$X_n = \begin{cases} |\Pi_T(P(T_n)) - u(T_n)|, & \text{if } P(T_n) \in \Phi, \\ \infty & \text{otherwise}, \end{cases} \tag{15}$$

$$Y_n = \begin{cases} \text{ycoord}\,(P(T_n)), & \text{if } P(T_n) \in \Phi, \\ \infty & \text{otherwise}. \end{cases} \tag{16}$$

Figure 7 illustrates the relative position of an association base station. Let us recall some terminology regarding general state space Markov processes.

Definition 7 Given a discrete time Markov process, $(\mathbf{Z}_n, n \geq 0)$, with $\mathbf{Z}_n = (X_n, Y_n)$ with a state space, $\mathcal{Z}$, the transition probability kernel, $\mathbb{P}(\mathbf{z}, Q)$ with $\mathbf{z} = (w, h)$ and Q a set in the Borel σ algebra of $\mathcal{Z}$ is defined as:

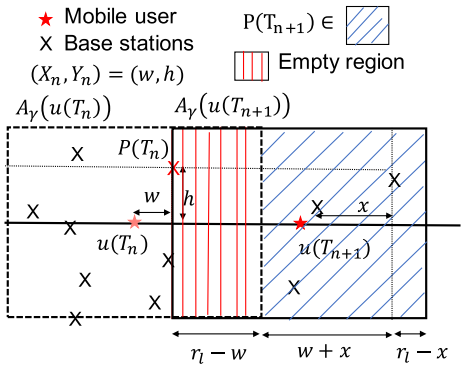

Figure 7: Relative position (X_n, Y_n) of the association base station $P(T_n)$ at time instant T_n

$$\mathbb{P}(\mathbf{z}, Q) = \mathbb{P}(\mathbf{z}_n \in Q | \mathbf{z}_{n-1} = \mathbf{z}). \tag{17}$$

If the transition kernel admits a density $f(\mathbf{z}, \bar{y})$ then,

$$\mathbb{P}(\mathbf{z}, Q) = \int_{\bar{y} \in Q} f(\mathbf{z}, \bar{y}) \mathrm{d}\bar{y}. \tag{18}$$

Leveraging the properties of the random geometry of base stations, i.e., the Poisson point process assumptions, we have the following result.

Theorem 1 *Under farthest greedy policy, if the handover support set is a rectangle $R_{(r_l,r_b)}(u)$ of length $2r_l$ and breadth $2r_b$,*

- *the x-coordinate of the relative position of the association base station $(X_n, n \in \mathbb{N})$ is a Markov process with state space $\mathcal{Z} = \{[-r_l, r_l], \infty\}$ and the transition kernel densities for x in set $Q(x) = [-r_l, x]$ are given by:*
- $$f(\infty, x) = \delta(x - r_l), \quad \mathbb{P}(w, \infty) = \mathrm{e}^{-2\lambda r_b(r_l + w)},$$
$$f(w, x) = \begin{cases} 2\lambda r_b \mathrm{e}^{-2\lambda r_b(r_l - x)}, & \text{if } -w \leq x \leq r_l, \\ 0 & \text{otherwise;} \end{cases} \tag{19}$$
- *the y-coordinate of the n^{th} relative position is given as:*
- $$Y_{n+1} = \begin{cases} Y, & \text{if } P(T_{n+1}) \in \Phi, \\ \infty & \text{otherwise}, \end{cases} \tag{20}$$

  where, $Y \sim U[-r_b, r_b]$ and the event that $\{P(T_{n+1}) \in \Phi\}$ depends only on X_n, i.e., given $X_n = w$, its probability is given by $1 - \mathrm{e}^{-2\lambda r_b(w + r_l)}$.

1 *Proof*

Recall that $(T_i)_{i=1}^{\infty}$ denotes handover times and $u(T_i)_{i=1}^{\infty}$ the locations of the mobile user at handover times. The n^{th} handover time, T_n is a stopping time, since the event $\{T_n = t\}$ for $t \geq 0$, depends only on $\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_n$. Then, conditionally on $T_n < \infty$ and $\mathbf{Z}_n = (w, h)$, we have that for all $m \geq 1$:

$$\begin{aligned} &\mathbb{P}(\mathbf{Z}_{n+m} = \mathbf{z} | \mathbf{Z}_i = \mathbf{z}_i \, \forall i < n, \mathbf{Z}_n = (w, h)) \\ &\quad = \mathbb{P}(\mathbf{Z}_{n+m} = \mathbf{z} | \mathbf{Z}_n = (w, h)). \end{aligned} \tag{21}$$

Recall that Z_{n+1} is the relative position of the association base station, $P(T_{n+1}) \in \Phi \cap R_{(r_l,r_b)}(u(T_{n+1}))$. Given our

greedy association policy, we have the following constraint that:

$$P(T_{n+1}) \in \Phi \cap \left[R_{(r_l,r_b)}(u(T_{n+1}))\backslash R_{(r_l,r_b)}(u(T_n))\right]. \tag{22}$$

This is due to the fact that a part of the region, $R_{(r_l,r_b)}(u(T_{n+1}))$ that is empty of base stations is based on its previous association, $P(T_n)$, as illustrated in Figure 7.

Further, note that one can determine the location $u(T_n)$ and its corresponding region $R_{(r_l,r_b)}(u(T_n)))$, based on $Z_n = (w, h)$ alone, in particular, only based on the x-coordinate of the previous state, i.e., X_n. Thus, the process tracking the x-coordinate of relative position of the association base station $(X_n, n \in \mathbb{N})$ satisfies the Markov property with state space, $\mathcal{Z} = \{[-r_l, r_l], \infty\}$. The y-coordinate Y_{n+1} is independent of Y_n and is uniformly distributed between $[-r_b, r_b]$, given that the user is connected as given in (20).

The discrete transition probability to the state "∞", $\mathbb{P}(w, \infty)$ is given by the probability that the set, $R_{(r_l,r_b)}(u(T_{n+1}))\backslash R_{(r_l,r_b)}(u(T_n))$, i.e., $R_2 \cup R_3$ as illustrated in Fig. 7 is empty. Given that the handover support set is empty of base stations, the first base station to enter will be at the edge. Thus, the Markov process always transitions from the state "∞" to a state "r_l", and therefore the transition probability kernel, $f(\infty, x)$ will be a Dirac delta function as given in (19).

The transition probability kernel for the continuous part is given by the product of the probability that the set R_2, the rectangle of length $(w + x)$ has at least one point and the set R_3, the rectangle of side length $(r_l - x)$ has no points, see Fig. 7:

$$\mathbb{P}(w, Q(x)) = \mathrm{e}^{-\lambda 2 r_b (r_l - x)}\left(1 - \mathrm{e}^{-\lambda 2 r_b (w+x)}\right), \quad -w \le x \le r_l. \tag{23}$$

Thus, the transition kernel density for the continuous and discrete parts of the state space $\mathcal{Z}$ is given by (19). □

Let us now focus on proving the existence and uniqueness of its stationary distribution. The Markov process $(X_n, n \ge 0)$ has the following properties:

- *Irreducibility* In the state space of the Markov process, the disconnected state, ∞, is an atom and the given Markov process is irreducible with respect to this atom.
- *Aperiodicity* There is no partition of the rectangle, $\mathcal{Z}$ in subsets $z_0, z_1, \ldots z_{k-1}$ such that the probability from transitioning from one subset to next set is one. Thus, the Markov process is aperiodic according to the definition of aperiodicity given in[31].

Given the properties of the Markov process and its transition kernel densities, we have the following result regarding its stationary distribution:

Theorem 2 *Under the rectangle-based association policy, the Markov process, $(X_n, n \ge 0)$ has a stationary distribution with discrete and continuous parts π_d and π_c given by its Laplace transforms:*

$$\begin{aligned}
&\mathcal{L}_{\pi_c}(s) \\
&= \frac{\left[2\lambda r_b s \mathrm{e}^{-sr_l} - 4\lambda^2 r_b^2 \mathrm{e}^{-4\lambda r_l r_b + sr_l}\right]\left(\frac{1+\mathcal{L}_{\pi_c}(0)}{2}\right)}{2\lambda r_b s - s^2 - 4\lambda^2 r_b^2 \mathrm{e}^{-4\lambda r_l r_b}} \\
&\quad + \frac{\left[4\lambda^2 r_b^2 \mathrm{e}^{-4\lambda r_l r_b - sr_l} - 2\lambda s r_b \mathrm{e}^{-4\lambda r_l r_b + sr_l}\right]\left(\frac{1-\mathcal{L}_{\pi_c}(0)}{2}\right)}{2\lambda r_b s - s^2 - 4\lambda^2 r_b^2 \mathrm{e}^{-4\lambda r_l r_b}}, \\
&\pi_d(r_l) = \pi_d(\infty) = q = \frac{1 - \mathcal{L}_{\pi_c}(0)}{2},
\end{aligned} \tag{24}$$

where,

$$\mathcal{L}_{\pi_c}(0) = \frac{2\lambda r_b\left(1 - \mathrm{e}^{2a_1 r_l}\right) + a_1\left(\mathrm{e}^{4\lambda r_l r_b} - \mathrm{e}^{2a_1 r_l}\right)}{2\lambda r_b\left(1 + \mathrm{e}^{2a_1 r_l}\right) - a_1\left(\mathrm{e}^{4\lambda r_l r_b} + \mathrm{e}^{2a_1 r_l}\right)}, \tag{25}$$

with $a_1 = \lambda r_b + \lambda r_b\sqrt{1 - 4\mathrm{e}^{-4\lambda r_l r_b}}$.

1 *Proof*

Let us consider the stationary distribution of this process in terms of π_d and π_c associated with the discrete and the continuous parts. Thus $\pi_c(x)$ represents the continuous part of the stationary distribution of the Markov process for $x \in [-r_l, r_l)$, $\pi_d(r_l)$ and $\pi_d(\infty)$ are the discrete probability masses associated with the states r and ∞. We have the following equations:

$$\begin{aligned}
&\pi_c(x) = 2\lambda r_b \mathrm{e}^{-2\lambda r_b (r_l - x)}\left(\int_{-x}^{r_l^-} \pi_c(w)\mathrm{d}w + \pi_d(r_l)\right), \\
&\pi_d(\infty) = \mathrm{e}^{-2\lambda r_b r_l}\int_{-r_l}^{r_l^-} \pi_c(w)\mathrm{e}^{-2\lambda r_b w}\mathrm{d}w + \pi_d(r_l)\mathrm{e}^{-4\lambda r_b r_l}, \\
&\pi_d(r_l) = \pi_d(\infty),
\end{aligned} \tag{26}$$

with $\pi_d(r_l) + \pi_d(\infty) + \int_{-r_l}^{r_l^-} \pi_c(w)\mathrm{d}w = 1$. We use Laplace transform techniques to solve the integral equations. The detailed evaluation is given in Appendix B. □

Let the relative position of the association base station conditioned on the availability of an accessible base station be denoted by $(\hat{X}_1, \hat{Y}_1)$. Given the Laplace transform of the stationary

distribution, we can evaluate various moments of the conditional random variable $\hat{X}_1$ and also study its asymptotics. The mean of the random variable $\hat{X}_1$ is given by:

$$\mathbb{E}[\hat{X}_1] = -\mathcal{L}'_{\pi_c}(0) = -\frac{(1-\mathrm{e}^{-4\lambda r_l r_b})[1-\mathcal{L}_{\pi_c}(0)] - 4\lambda r_l^2 \mathrm{e}^{-4\lambda r_l r_b}}{4\lambda r_l \mathrm{e}^{-4\lambda r_l r_b}}. \tag{27}$$

In the limit, as the density of base station λ goes to infinity, we have that

$$\lim_{\lambda\to\infty} \mathcal{L}_{\pi_c}(0) = 1.$$

Thus, we have that the stationary distribution of the states r_l and ∞ goes to zero, i.e., $\pi_d(r_l) = \pi_d(\infty) = 0$, and for the continuous part we have:

$$\lim_{\lambda\to\infty} \pi_c(x) = \delta(x - r_l),$$

and, in the limit, the expectation is $\mathbb{E}[\hat{X}_1] = r_l$.

4.4 Evaluation of the Rate of Handovers

Given the Markov process characterization of the greedy association policy, one can evaluate the rate of handovers seen by a mobile user as stated in the following theorem:

Theorem 3 *Consider a mobile user moving along a straight line at constant velocity v in a Poisson cellular network with base station intensity, λ. Under the rectangle-based association policy with accessible base stations defined by $R_{(r_l,r_b)}(u)$, the rate of handovers seen by the mobile user, ρ is given by*

$$\rho = \frac{2\lambda r_b v}{2\lambda r_l r_b + q + 2\lambda r_b(1-2q)(-\mathcal{L}'_{\pi_c}(0))}, \tag{28}$$

where, $q = \pi_d(r_l) = \frac{1-\mathcal{L}_{\pi_c}(0)}{2}$ and $\mathcal{L}'_{\pi_c}(0)$ is given by (27).

1 Proof

Recall that the sequence $(T_i)_{i=1}^{\infty}$ denotes the handover times and that at any time, T_n, X_n is the relative x-coordinate of the associated base station. Then, we have

$$T_1 = \begin{cases} \frac{\hat{X}_1 + r_l}{v} & \text{w.p. } 1-2q, \\ \frac{2r_l}{v} & \text{w.p. } q, \\ G & \text{w.p. } q, \end{cases} \tag{29}$$

where, q is given in (24) and $G \sim \exp(2\lambda r_b v)$ is the time the rectangle $R_{(r_l,r_b)}(u)$ is empty of points. The rate of handovers is then given by:

$$\rho = \frac{1}{\mathbb{E}^0[T_1]}, \tag{30}$$

where $\mathbb{E}^0$ denotes the Palm expectation which is given by:

$$\mathbb{E}^0[T_1] = (1-2q)\mathbb{E}^0\left[\frac{\hat{X}_1 + r_l}{v}\right] + \frac{2qr_l}{v} + \frac{q}{2\lambda r_b v} = \frac{r_l}{v} + \frac{q}{2\lambda r_b v} + \frac{1-2q}{v}(-\mathcal{L}'_{\pi_c}(0)), \tag{31}$$

since, $\mathbb{E}^0[\hat{X}_1] = -\mathcal{L}'_{\pi_c}(0)$, which is given by (27). □

Corollary 1 Under the rectangle-based association policy, we have that

$$\lim_{\lambda\to\infty} \rho = \frac{v}{2r_l}. \tag{32}$$

1 Proof

As the density of base stations, λ, goes to infinity, we have that $\lim_{\lambda\to\infty} \mathcal{L}_{\pi_c}(0) = 1$, and $\mathbb{E}[\hat{X}_1] = r_l$. Thus, we have that $q = 0$ and the rate of handovers is a constant $\frac{v}{2r_l}$. □

4.5 Evaluation of Effective Throughput

Given the steady-state formulation of the Markov process, the expectation of the rate can be evaluated using the inversion formula[32] as:

$$\mathcal{S} = \mathbb{E}[S(0)] = \frac{\mathbb{E}^0\left[\int_0^{(T_1-\delta)^+} S(0)o\,\theta_t \mathrm{d}t\right] - c}{\mathbb{E}^0[T_1]}, \tag{33}$$

where θ_t is the shift operator[32]. For mathematical simplicity, we consider that the user is served at all times and we take the effect of handovers into account by introducing a cost term as follows:

$$\hat{\mathcal{S}} = \frac{\mathbb{E}^0\left[\int_0^{T_1} S(0)o\,\theta_t \mathrm{d}t\right]}{\mathbb{E}^0[T_1]}(1-\rho\delta) - \rho c, \tag{34}$$

where $\rho = \frac{1}{\mathbb{E}^0[T_1]}$ is the rate of handovers and c is the fixed penalty for each handover. Note that the handover cost $\delta\rho$ is a unit-less quantity used to quantify the fraction of interruption time along the user trajectory.

First, let us consider the distance process

$$D(t) = \sqrt{(vt - X_n)^2 + Y_n^2}, \quad \text{for } T_n \le t < T_{n+1}, \tag{35}$$

if it is connected and ∞ otherwise. Let the relative position of the association base station conditioned on the availability of an accessible base station be denoted by $(\hat{X}_1, \hat{Y}_1)$. Given the distribution of random variable T_1 as in (29), and $q = \pi_d(\infty)$, the expectation is evaluated as follows:

$$\begin{aligned}&\mathbb{E}^0\left[\int_0^{T_1} S(0) o\, \theta_t \mathrm{d}t\right] \\ &= q\mathbb{E}^0\left[\int_0^{2r_l/v} h(r_l, \hat{Y}_1, t)\mathrm{d}t\right] \\ &\quad + (1-2q)\mathbb{E}^0\left[\int_0^{(\hat{X}_1+r_l)/v} g(\hat{X}_1, \hat{Y}_1, t)\mathrm{d}t\right],\end{aligned} \tag{36}$$

where

$$g(\hat{X}_1, \hat{Y}_1, t) = \omega \ln\left(1 + \frac{\xi \times l(\sqrt{(vt - \hat{X}_1)^2 + \hat{Y}_1^2})}{\sigma^2}\right),$$

and

$$h(r_l, \hat{Y}_1, t) = \omega \ln\left(1 + \frac{\xi \times l(\sqrt{(vt - r_l)^2 + \hat{Y}_1^2})}{\sigma^2}\right).$$

Special Case We evaluate the effective throughput to get a closed-form analytical expression by considering a specific value for the fitting parameter $\beta = 2$ for the bounded path loss function in (1). Approximating $\ln(1+x) \sim \ln(x)$, for large x, we have that:

$$\begin{aligned}g(\hat{X}_1, \hat{Y}_1, t) &= \omega\Big(\ln(\xi/\sigma^2) - k\Big[(vt - \hat{X}_1)^2 + \hat{Y}_1^2\Big]\Big), \\ h(r_l, \hat{Y}_1, t) &= \omega\Big(\ln(\xi/\sigma^2) - k\Big[(vt - r_l)^2 + \hat{Y}_1^2\Big]\Big).\end{aligned} \tag{37}$$

Let $m = \ln(\xi/\sigma^2)$. Now, given $\hat{Y}_1 \sim U[-r_b, r_b]$, one can evaluate the expectation in (36), which gives:

$$\begin{aligned}&\mathbb{E}[S(0)] \\ &= \frac{q\omega}{\mathbb{E}^0[T_1]}\left[\frac{6r_l m - 2\kappa r_l^3 - 2r_l r_b^2 \kappa}{3v}\right] \\ &\quad + \frac{(1-2q)\omega}{\mathbb{E}^0[T_1]}\left[\frac{3m\mathbb{E}^0[X_1] - \kappa\mathbb{E}^0[X_1^3] + 3mr_l - \kappa r_l^3 - \kappa r_l r_b^2}{3v}\right],\end{aligned} \tag{38}$$

where $q = \pi_d(\infty)$ is given in (24), and $\mathbb{E}^0[T_1]$ is given in (31).

4.6 Optimal Farthest Greedy Policy

The problem of finding the optimal policy is now posed as a parametric optimization problem as follows:

Problem 2 The problem of finding an optimal sequence of association base stations and handover times under rectangle-based farthest greedy association policy consists in finding the optimal parameters (r_l, r_b) of the set $R_{(r_l,r_b)}(u)$ that maximize the effective throughput $\mathcal{S}((r_l, r_b))$.

$$\mathcal{S}^* = \max_{r_l, r_b}\{\mathcal{S}((r_l, r_b))\}. \tag{39}$$

Associating with a base station at a greater distance decreases the handovers but may result in low rate. Thus, we focus on optimizing the stationary effective throughput with respect to the constraint on the greediness, i.e., the size of the handover support set.

If we consider a square handover support set of side length $r = r_l = r_b$, we have that in the limit r going to zero and r going to ∞, the effective throughput in (36) goes to zero. Thus, a local or global optimum value must exist for the effective throughput. Note that it is challenging to give a closed-form expression for the optimal parameters r_l^*, r_b^* which maximizes effective throughput, thus we evaluate the optimal parameter values with the help of simulations in Sect. 5

5 Numerical Results and Simulations

In this section, we study the effectiveness of our proposed association policies with the help of simulation. We compare their performance and assess their improvements with respect to the closest base station association policy. We then study the performance sensitivity of these policies to the velocity, cost of handovers and handover delay. Unless otherwise stated, we consider a realization of a Poisson point process of intensity λ and the parameters of Table 3. In particular, we focus on the following questions:

Table 3 Table of simulation parameters

Parameter	Value
Power ξ	1 W
Bandwidth ω	30 MHz
Velocity v	20 m/s
Handover delay δ	20 ms
Handover cost c	100 Kbits

- What is the performance and complexity trade-off for the optimal forward association policies with respect to the size of the handover support set $A_\gamma(u)$?
- What is the trade-off between the farthest distance and the height of the handover support set, i.e., the length and breadth in rectangular support set with respect to performance under greedy policies?
- What are the performance improvements of the optimal forward policy over the optimal greedy policy under various environments?
- When do the performance metrics of both policies have low discrepancy?

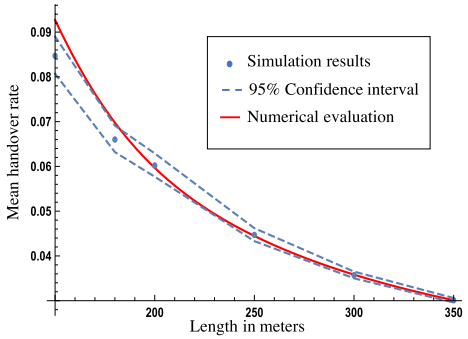

Figure 8: Comparison of the numerical results with simulation for various lengths of the square region and $\lambda = 40$ per km ^{2}.

5.1 Optimal Forward Association Policy

In this section, we consider handover support set $A_\gamma(u)$ with the tagged user traveling a distance of 10 km to destination (10, 0) on the positive x-axis. We build the directed acyclic graph as in Definition 3 and solve the problem of the maximum weight path using Dijkstra's algorithm.

Problem 1 consists of finding the association sequence $\mathbf{p}^* \in \mathcal{P}_f$ that results in a maximum reward, i.e., the total volume of data delivered for fixed handover support set $A_\gamma(u)$. Note that increasing the size of the handover support set leads to more feasible association sequences, i.e., more nodes in the graph and hence higher complexity. We notice that beyond a certain size of the handover support set, increasing the set does not improve the optimization. Thus for our simulation study, we consider one fixed size, $r_l = 400$ m for the square region, the value of which depends on the decay of the path loss function considered.

5.2 Farthest Greedy Association Policy

For rectangular-based farthest greedy policies, we have a closed-form expression for the effective throughput and the rate of handovers. In this subsection, we first compare the numerical results with the simulation. We then consider the dependence of the optimal parameters on the intensity of base stations λ.

5.2.1 Evaluation of the Rate of Handovers

We numerically evaluated the closed-form expression for the mean rate of handovers as given in (30), for the intensity of base stations, $\lambda = 40$ per km ^{2}. We consider a square handover support set, i.e., $r_l = r_b$ and compare the numerical results with that of the simulation. The simulation results agree with the numerical evaluation for various values of the side of the square as illustrated in Fig. 8.

5.2.2 Evaluation of Effective Throughput

We focus on the trade-off between the length and breadth of the region with respect to performance for different (1) intensities of base stations, and (2) path loss functions. While the cost of handover depends on how far in the direction of motion the user can associate with, i.e., the length of the region, the rate is seen by the user depends on how close the base station is to the trajectory, i.e., the breadth of the region.

Intensity of Base Stations We numerically evaluate the effective throughput for various values of the parameters r_l, r_b for two intensities of base stations $\lambda = 20$ per km ^{2} and $\lambda = 200$ per km ^{2}. Figure 9 illustrates the surface plots of the effective throughput for the path loss function given in (1), with $\beta = 2$.

We have two insights: (1) when increasing λ, the probability of finding a base station at the forward edge of the rectangular region increases, thus the optimal size is smaller for higher densities of base stations, and (2) given that a base station that is closer to the tagged user's trajectory leads to a higher rate, the shape of the optimal regions is an elongated rectangle, i.e., the length is higher than the breadth.

Path Loss Functions We evaluate the effective throughput with the help of simulation for the

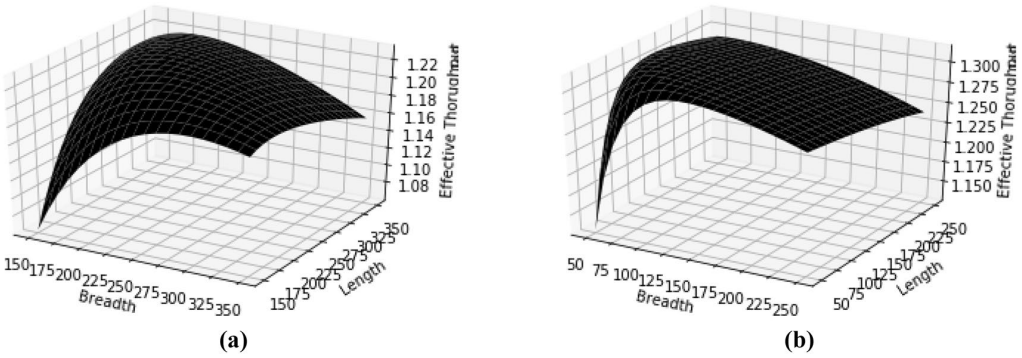

Figure 9: Surface plots of effective throughput for two intensities of base stations. **a** For $\lambda = 20$, the optimal parameters $(r_l^*, r_b^*) = (286.8, 171)$ m and **b** for $\lambda = 200$, the optimal parameters $(r_l^*, r_b^*) = (176.3, 113.1)$ m.

path loss function given in (1) with two different fitting parameters $\beta = 2, 1.5$,

$$l_1(x) = \exp(-\kappa x^{3/2}), l_2(x) = \exp(-\kappa x^2), \tag{40}$$

for a given intensity of base stations $\lambda = 20$ per km ^{2}. In Fig. 10, we give the surface plots of the effective throughput. The optimal size of the region is smaller in the case where there is faster decay in the path loss function. Similarly, the shape of the optimal regions is an elongated rectangle, i.e., the length is higher than the breadth.

5.2.3 Parametric Optimization under Constraint

We now consider the parametric optimization problem under the additional constraint that the rectangular handover support set has a fixed area. There is a certain amount of complexity/overhead involved in gaining knowledge about the base station locations. Thus we limit the complexity by considering a fixed area of the rectangular handover support set. In this setting, we evaluate the effective throughput with the help of simulations for various intensities of the base station and an area of $4r_l r_b = 10^4$ m ^{2}, as illustrated in Fig. 11, under the area constraint. The optimal rectangular shape is narrower for a higher intensity of base stations than before.

5.3 Comparison of Performance of Optimal Forward Association and Optimal Farthest Greedy Policy

In this subsection, we compare the optimal farthest greedy policy of Sect. 4 with the dynamic programming based policy of Sect. 3 by considering a common realization of a Poisson point process on which they are jointly applied. We consider rectangular handover support sets $R_{(r_l, r_b)}(u)$ with the optimal parameters for the greedy policies that are evaluated as in the previous section. We consider the total volume of data delivered as the common performance metric. To account for the variability of the Poisson point

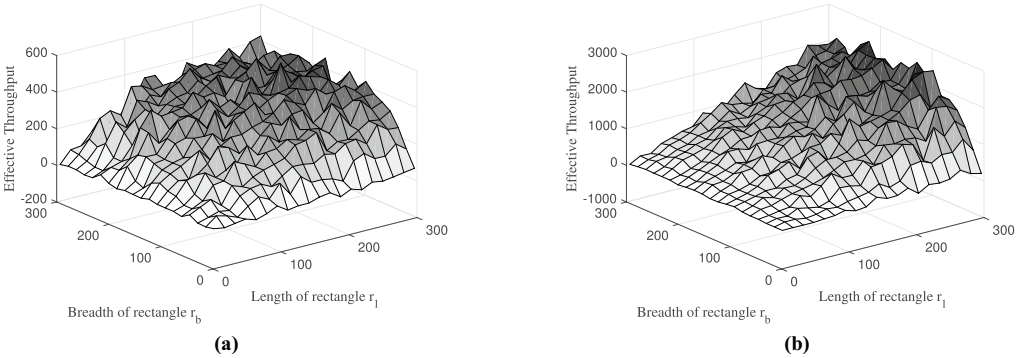

Figure 10: Surface plots of effective throughput for two path loss functions. **a** For $l_1(x)$, the optimal parameters $r_l^* \in [260, 280]$ m, $r_b^* \in [180, 200]$ m and **b** for $l_2(x)$, the optimal parameters $r_l^* \in [260, 280]$ m, $r_b^* \in [180, 200]$ m

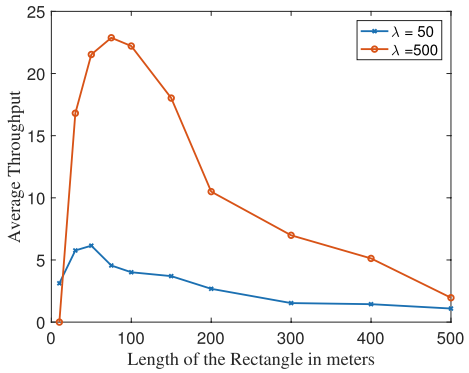

Figure 11: Comparing the effective throughput for various intensities of base stations under the constraints of fixed area.

process, we average the total volume of data delivered for many realizations of the point process.

The global picture is that the optimal forward association policy has roughly the same performance as the greedy policy under the considered parameters. We evaluated the performance for intensity of base stations $\lambda = 20, 80, 200$ per km ^{2}. The percentage of increase in the total volume of data delivered by the dynamic programming approach compared to the greedy policy is in the range of 1–5%. The improvement tends to be 0 when increasing (1) the handover delay, (2) the cost of each handover, or (3) the velocity of the user, as illustrated in Fig. 12.

The optimal greedy association policy gives comparable results (within 1–5 mobile users moving at a velocity greater than 10 m/s and for realistic values of handover delay ($d > 100$ ms) and handover cost ($c > 100$ kbits) across all density of base stations ($\lambda = 20$–200 per km ^{2}).

Further, we compare the performance of our policies with the traditional closest base station policy. We have that the improvement is larger for the high cost of handovers and handover delays in dense networks as illustrated in Fig. 12.

5.4 Comparison of Performance of Optimal Forward Association Policy with 3GPP Standard

In this subsection, we compare the optimal farthest greedy policy with the performance of mobile users following the present standard. In LTE, there are eight types of handover event entry conditions (see[33] section 5.5.4). Mobile users periodically do measurements for the neighboring base stations using the reference signal. Handovers are then triggered by the mobile user based on such measurements of RSRP (reference signal received power) values of the serving and the target base stations. The final decision to go ahead with a handover is made by the serving base station. Most of the handovers are triggered when the signal from the target base station is better than an offset from that of the serving base station.

Handovers can be governed not only by signal strength but also by signal quality. In this paper, we use SNR as a measure of signal quality. Even though it is not defined in the 3GPP specs, it is defined by vendors and used to calculate channel quality index (CQI).

We consider rectangular handover support sets $R_{(r_l,r_b)}(u)$, where every 10 ms the mobile user measures for SNR from all the base stations within its handover support set i.e., neighboring cells. We evaluate the effective throughput by penalizing the handovers as before. We compare this to our proposed model, where mobile user associates with a base station farthest in the direction of its motion.

To better reflect the mobile user's motion in real-life scenarios, we consider a variant of the

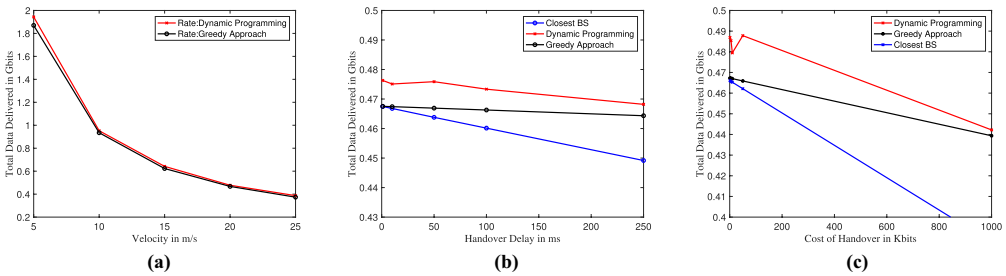

Figure 12: Total data delivered vs **a** velocity ($\lambda = 20$); **b** handover delay ($\lambda = 80$); **c** cost of handover ($\lambda = 80$).

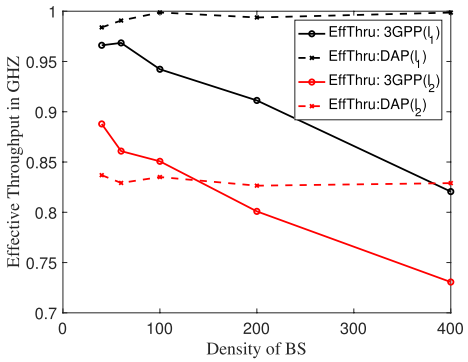

Figure 13: Effective throughput vs density of base stations for path loss function $l_1(r) = \frac{1}{(1+0.01r)^4}$, path loss function $l_2(r) = \frac{1}{(1+r)^4}$.

random way point model,[34], where a user moves at a constant velocity, v meters per sec, but every t_0 secs, with probability p_0 selects a new independent direction (angle) uniformly distributed in $(-\theta_0, \theta_0)$.

We also consider the effect of channel fading, by considering SNR at a location u when associated with BS at location p_i as follows:

$$\text{SNR}\ (u) = \frac{H\xi l(||p_i - u||)}{\sigma^2}, \tag{41}$$

where H is a fading random variable which is exponential with unit mean.

First, for a given fixed size of the rectangular regions of $(r_l, r_b) = (500, 200)$ m, we evaluate the effective throughput for both cases for increasing density of base stations. Figure 13 illustrates the decrease in the effective throughput of the standard with the increasing density of base stations due to the increase in the number of handovers. Further, the density of base stations after which our policy performs better than the standard depends on various parameters like the size of the region and the decay of the path loss function as illustrated in Fig. 13.

6 Conclusion

In dense networks, the increased intensity of handovers has a negative impact on mobile users' performance. In this paper, we proposed strategies for association that achieve a reasonable throughput–handover trade-off. The first strategy takes the user's trajectory and the locations of all base stations into account and determines an optimal set of association base stations and handover times. This is achieved by finding the maximum weight path in a directed acyclic graph.

The second strategy, to use limited knowledge about the base station locations, is a direction-dependent greedy association policy that results in a reduced number of handovers. We give a closed-form expression for the stationary effective throughput seen by the user under this policy based on a Markovian formulation. This is in turn used to derive the optimal farthest greedy association policy.

We validate the effectiveness of this optimized greedy policy by comparing it with dynamic programming in various scenarios. The percentage gains achieved by the dynamic programming policy are within 1–5% from the greedy policies for the considered parameters.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements

The research in this paper was supported in part by the National Science Foundation under Grant numbers NSF-CCF-1514275, NSF-CNS (1731658) and an award from the Simons Foundation (197982), all to the University of Texas at Austin.

Appendix

Optimal Association Policy

Problem 3 The optimal handover problem consists in finding a sequence of base stations and handover times that results in the maximum volume of data, i.e.,

$$r^* = \max_{\mathbf{p},\mathbf{t}} \{r(\mathbf{p}, \mathbf{t}) : \mathbf{p} \in \mathcal{P}, \mathbf{t} \in \mathcal{T}_\mathbf{p}\}, \tag{42}$$

where $\mathcal{P}, \mathcal{T}_\mathbf{p}$ denote the set of all possible association sequences and associated handover times, respectively.

For each handover from base station $p_i \to p_{i+1}$, we have a continuum of possible handover times with some constraints that we will not specify. Thus, the above problem is a mixed optimization with discrete and continuous variables can be very challenging to solve. Further, the handover time t_i is functionally dependent on the entire association sequence.

Let us discretize the time with step size δ and consider a decision tree with origin o as the parent node and feasible association base stations and disconnected states at each time step as a new layer in a decision tree. If m denotes the total number of time steps, then the number of nodes in such a tree is in the order of $(2\lambda(\pi r^2))^m$. Further, the combinatorial optimization problem reduces to finding a maximum reward path in this decision tree, which has a complexity of $O((2\lambda(\pi r^2))^m)$[30].

Proof of Theorem 2

Let us consider the stationary distribution of this process in terms of π_d and π_c associated with the discrete and the continuous parts. Thus, $\pi_c(x)$ represents the continuous part of the stationary distribution of the Markov process for $x \in [-r_l, r_l)$, $\pi_d(r_l)$ and $\pi_d(\infty)$ are the discrete probability masses associated with the states r_l and ∞.We have the following equations:

$$\begin{aligned}
&\pi_c(x) = 2\lambda r_b \mathrm{e}^{-2\lambda r_b(r_l - x)}\left(\int_{-x}^{r_l^-} \pi_c(w)\mathrm{d}w + \pi_d(r_l)\right),\\
&\pi_d(\infty) = \mathrm{e}^{-2\lambda r_b r_l}\int_{-r_l}^{r_l^-} \pi_c(w)\mathrm{e}^{-2\lambda r_b w}\mathrm{d}w + \pi_d(r_l)\mathrm{e}^{-4\lambda r_b r_l},\\
&\pi_d(r_l) = \pi_d(\infty),
\end{aligned} \tag{43}$$

with $\pi_d(r_l) + \pi_d(\infty) + \int_{-r_l}^{r_l^-} \pi_c(w)\mathrm{d}w = 1$. We use Laplace transform techniques to solve the integral equations. First, we multiply on both sides of the first equation in (43) with e^{-sx} and integrate from $-r_l$ to r_l, we get

$$\begin{aligned}
\int_{-r_l}^{r_l} \pi_c(x)\mathrm{e}^{-sx}\mathrm{d}x = \frac{2\lambda r_b}{(2\lambda r_b - s)}\Big[&\mathrm{e}^{-sr_l} - q\mathrm{e}^{-4\lambda r_b r_l + sr_l}\\
&+\mathrm{e}^{-sr_l}\int_{-r_l}^{r_l} \pi_c(z)dz - \mathrm{e}^{-2\lambda r_b r_l}\int_{-r_l}^{r_l} \pi_c(z)\mathrm{e}^{-(2\lambda r_b - s)z}dz\Big]
\end{aligned} \tag{44}$$

Let $\mathcal{L}_{\pi_c}(s) = \int_{-r_l}^{r_l} \pi_c(x)\mathrm{e}^{-sx}\mathrm{d}x$, then we have

$$\begin{aligned}
\mathcal{L}_{\pi_c}(s) = \frac{2\lambda r_b}{(2\lambda r_b - s)}\Big[&\mathrm{e}^{-sr}\mathcal{L}_{\pi_c}(0)\\
&-\mathrm{e}^{-2\lambda r_b r_l}\mathcal{L}_{\pi_c}(2\lambda r_b - s)\\
&+q\mathrm{e}^{-sr_l} - q\mathrm{e}^{-4\lambda r_b r_l + sr_l}\Big],
\end{aligned} \tag{45}$$

where $q = \pi_d(r) = \frac{1-\mathcal{L}_{\pi_c}(0)}{2}$.

Let $\alpha(s)$ and $\beta(s)$ be two functions such that

$$\begin{aligned}
\mathcal{L}_{\pi_c}(s) &= \alpha(s) + \beta(s)\mathcal{L}_{\pi_c}(2\lambda r_b - s)\\
&= \alpha(s) + \beta(s)[\alpha(2\lambda r_b - s)\\
&\quad + \beta(2\lambda r - s)\mathcal{L}_{\pi_c}(s)]\\
&= \frac{\alpha(s) + \beta(s)\alpha(2\lambda r_b - s)}{1 - \beta(s)\beta(2\lambda r_b - s)}.
\end{aligned} \tag{46}$$

From (45) we have:

$$\begin{aligned}
&\alpha(s) = \frac{2\lambda r_b}{(2\lambda r_b - s)}\left[\mathrm{e}^{-sr_l}\mathcal{L}_{\pi_c}(0) + q\mathrm{e}^{-sr_l} - q\mathrm{e}^{-4\lambda r_b r_l + sr_l}\right],\\
&\beta(s) = -\frac{2\lambda r_b \mathrm{e}^{-2\lambda r_b r_l}}{(2\lambda r_b - s)}.
\end{aligned} \tag{47}$$

Since the range of the random variable with probability distribution π_c is finite, the Laplace transform cannot be degenerate. Therefore, the zeros of the denominator in (46), should match the zeros of the numerator. Thus, the zeros a_1, a_2 of the denominator are given by solving $1 = \beta(s)\beta(2\lambda r_b - s)$ which gives the following quadratic equation

$$s^2 - 2\lambda r_b s + 4\lambda^2 r_b^2 \mathrm{e}^{-4\lambda r_b r_l} = 0. \tag{48}$$

Thus, the roots of the denominator are $a_1, a_2 = \lambda r_b \pm \lambda r_b\sqrt{1 - 4\mathrm{e}^{-4\lambda r_b r_l}}$. Note, that the roots are imaginary for $4\lambda r_b r_l < 1.38$. Thus, the mean number of base stations$(\lambda 4 r_b r_l)$ in the rectangular region of side lengths $2r_l, 2r_b$ considered has to be greater than 1.38 for real roots.

Now, let us denote the numerator in (46) as $f(s)$, then we have that $f(a_1) = f(a_2) = 0$, using which we will solve for $\mathcal{L}(0)$. The function $f(s)$ is given as:

$$\begin{aligned}
f(s) = &\left[2\lambda r_b s\mathrm{e}^{-sr_l} - 4\lambda^2 r_b^2 \mathrm{e}^{-4\lambda r_b r_l + sr_l}\right]\left(\frac{1 + \mathcal{L}_{\pi_c}(0)}{2}\right)\\
&+\left[4\lambda^2 r_b^2 \mathrm{e}^{-4\lambda r_b r_l - sr_l} - 2\lambda s r_b \mathrm{e}^{-4\lambda r_b r_l + sr_l}\right]\left(\frac{1 - \mathcal{L}_{\pi_c}(0)}{2}\right).
\end{aligned} \tag{49}$$

Then, $f(\lambda r_b + \lambda r_b\sqrt{1 - 4\mathrm{e}^{-4\lambda r_b r_l}}) = 0$, yields (50).

$$\mathcal{L}_{\pi_c}(0) = \frac{2\lambda r_b\left(1 - \mathrm{e}^{2a_1 r_l}\right) + a_1\left(\mathrm{e}^{4\lambda r_l r_b} - \mathrm{e}^{2a_1 r_l}\right)}{2\lambda r_b\left(1 + \mathrm{e}^{2a_1 r_l}\right) - a_1\left(\mathrm{e}^{4\lambda r_l r_b} + \mathrm{e}^{2a_1 r_l}\right)}, \tag{50}$$

where, $a_1 = \lambda r + \lambda r\sqrt{1 - 4\mathrm{e}^{-4\lambda r^2}}$.

Similarly for the root $a_2 = \lambda r - \lambda r\sqrt{1-4e^{-4\lambda r^2}}$, we get:

$$\mathcal{L}_{\pi_c}(0) = \frac{2\lambda r_b\left(1-e^{2a_2r_l}\right)+a_2\left(e^{4\lambda r_l r_b}-e^{2a_2r_l}\right)}{2\lambda r_b\left(1+e^{2a_2r_l}\right)-a_2\left(e^{4\lambda r_l r_b}+e^{2a_2r_l}\right)}, \tag{51}$$

Let us prove that both the above functions are same. Since a_1 and a_2 are roots of the quadratic equation given in (48), we have that:

$$\begin{aligned} a_1 + a_2 &= 2\lambda r_b, \\ a_1 a_2 &= 4\lambda^2 r_b^2 e^{-4\lambda r_b r_l}. \end{aligned} \tag{52}$$

Let us substitute $2\lambda r_b$ in (50) with $\frac{a_1 a_2 e^{4\lambda r_b r_l}}{2\lambda r_b}$:

$$\begin{aligned} \mathcal{L}_{\pi_c}(0) &= \frac{a_2 e^{4\lambda r_b r_l}\left(1-e^{2a_1r_l}\right)+2\lambda r_b\left(e^{4\lambda r_l r_b}-e^{2a_1r_l}\right)}{a_2 e^{4\lambda r_b r_l}\left(1+e^{2a_1r_l}\right)-2\lambda r_b\left(e^{4\lambda r_l r_b}+e^{2a_1r_l}\right)}, \\ &= \frac{a_2\left(1-e^{4\lambda r_l r_b e^{-2a_2r_l}}\right)+2\lambda r_b\left(1-e^{-2a_2r_l}\right)}{a_2\left(1+e^{4\lambda r_l r_b e^{-2a_2r_l}}\right)-2\lambda r_b\left(1+e^{-2a_2r_l}\right)}, \end{aligned} \tag{53}$$

where the last equality is obtained by substituting $a_1 = 2\lambda r_b - a_2$. Now, we get (51) by multiplying the numerator and the denominator of the above equation by $-e^{2a_2r_l}$.

Received: 15 December 2019 Accepted: 25 March 2020
Published online: 16 April 2020

References

1. Cisco Visual Networking Index (2015) Global mobile data traffic forecast update, 2014–2019, White Paper, February, vol 1
2. Andrews JG, Buzzi S, Choi W, Hanly SV, Lozano A, Soong AC, Zhang JC (2014) What will 5G be? IEEE J Sel Areas Commun 32(6):1065–1082
3. Tso FP, Teng J, Jia W, Xuan D (2010) Mobility: a double-edged sword for HSPA networks: a large-scale test on Hong Kong mobile HSPA networks. In: Proceedings of the eleventh ACM international symposium on mobile ad hoc networking and computing. ACM, pp 81–90
4. Lunden P, Aijanen J, Aho K, Ristaniemi T (2008) Performance of VoIP over HSDPA in mobility scenarios. In: Vehicular technology conference, 2008. VTC Spring 2008. IEEE. IEEE, pp 2046–2050
5. Sinclair N, Harle D, Glover IA, Atkinson RC (2012) A kernel methods approach to reducing handover occurrences within LTE. In: Wireless conference (European Wireless), 2012 18th European. VDE, pp 1–8
6. Kong P-Y, Sluzek A (2015) Average packet delay analysis for a mobile user in a two-tier heterogeneous cellular network. IEEE Syst J
7. Zhang H, Ma W, Li W, Zheng W, Wen X, Jiang C (2011) Signalling cost evaluation of handover management schemes in LTE-advanced femtocell. In: Vehicular technology conference (VTC Spring), 2011 IEEE 73rd. IEEE, pp 1–5
8. Zhang H, Jiang C, Cheng J, Leung VC (2015) Cooperative interference mitigation and handover management for heterogeneous cloud small cell networks. IEEE Wirel Commun 22(3):92–99
9. Ulukus S, Pollini GP (1998) Handover delay in cellular wireless systems. In: Communications, 1998. ICC 98. Conference record. 1998 IEEE international conference on, vol 3. IEEE, pp 1370–1374
10. Rengarajan B, De Veciana G (2011) Practical adaptive user association policies for wireless systems with dynamic interference. IEEE/ACM Trans Netw 19(6):1690–1703
11. Ye Q, Rong B, Chen Y, Al-Shalash M, Caramanis C, Andrews JG (2013) User association for load balancing in heterogeneous cellular networks. IEEE Trans Wirel Commun 12(6):2706–2716
12. Das S, Viswanathan H, Rittenhouse G (2003) Dynamic load balancing through coordinated scheduling in packet data systems. In: INFOCOM 2003. Twenty-second annual joint conference of the IEEE computer and communications. IEEE Societies, vol 1. IEEE, pp 786–796
13. Bonald T, Borst S, Proutiere A (2005) Inter-cell scheduling in wireless data networks. In: Wireless conference 2005-next generation wireless and mobile communications and services (European wireless), 11th European. VDE, pp 1–7
14. Sang A, Wang X, Madihian M, Gitlin RD (2008) Coordinated load balancing, handoff/cell-site selection, and scheduling in multi-cell packet data systems. Wirel Netw 14(1):103–120
15. Ramjee TBLELR, Bu T, Li L (2006) Generalized proportional fair scheduling in third generation wireless data networks. In: IEEE INFOCOM, pp 1–12
16. Kim H, de Veciana G, Yang X, Venkatachalam M (2010) Alpha-optimal user association and cell load balancing in wireless networks. In: INFOCOM, 2010 proceedings IEEE. IEEE, pp 1–5
17. Son K, Chong S, De Veciana G (2009) Dynamic association for load balancing and interference avoidance in multi-cell networks. IEEE Trans Wirel Commun 8:7
18. Son K, Kim H, Yi Y, Krishnamachari B (2011) Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks. IEEE J Sel Areas Commun 29(8):1525–1536
19. Baccelli F, Blaszczyszyn B, Muhlethaler P (2006) An Aloha protocol for multihop mobile wireless networks. IEEE Trans Inf Theory 52(2):421–436
20. Baccelli F, Blaszczyszyn B (2009) Spatial modeling of wireless communications–a stochastic geometry approach. Foundations and trends in networking. NOW Publishers, Delft

21. Lin X, Ganti RK, Fleming PJ, Andrews JG (2013) Towards understanding the fundamentals of mobility in cellular networks. IEEE Trans Wirel Commun 12(4):1686–1698
22. Bao W, Liang B (2015) Stochastic geometric analysis of user mobility in heterogeneous wireless networks. IEEE J Sel Areas Commun 33(10):2212–2225
23. Arshad R, ElSawy H, Sorour S, Al-Naffouri TY, Alouini M-S (2016) Handover management in dense cellular networks: a stochastic geometry approach. In: Communications (ICC), 2016 IEEE international conference on. IEEE, pp 1–7
24. Arshad R, ElSawy H, Sorour S, Al-Naffouri TY, Alouini M-S (2016) Cooperative handover management in dense cellular networks. In: Global communications conference (GLOBECOM), 2016 IEEE. IEEE, pp 1–6
25. AlAmmouri A, Andrews JG, Baccelli F (2017) Sinr and throughput of dense cellular networks with stretched exponential path loss. IEEE Trans Wirel Commun 17(2):1147–1160
26. Dimou K, Wang M, Yang Y, Kazmi M, Larmo A, Pettersson J, Muller W, Timner Y (2009) Handover within 3GPP LTE: design principles and performance. In: Vehicular technology conference fall (VTC 2009-Fall), 2009 IEEE 70th. IEEE, pp 1–5
27. Mahmoodi T, Seetharaman S (2014) On using a SDN-based control plane in 5G mobile networks. In: Wireless world research forum, 32nd Meeting. Citeseer
28. Madadi P, Baccelli F, de Veciana G (2016) On temporal variations in mobile user SNR with applications to perceived QoS. In: Modeling and optimization in mobile, ad hoc, and wireless networks (WiOpt), 2016 14th international symposium on. IEEE, pp 1–8
29. Ahuja RK, Magnanti TL, Orlin JB (1988) Network flows
30. Cormen TH, Leiserson CE, Rivest RL, Stein C (2001) Introduction to algorithms. MIT press, Cambridge, p 819
31. Roberts GO, Rosenthal JS et al (2004) General state space Markov chains and MCMC algorithms. Probab Surv 1:20–71
32. Baccelli F, Brémaud P (2013) Elements of queueing theory: Palm Martingale calculus and stochastic recurrences. Springer, Berlin, p 26
33. T. 36.331 Radio resource control; protocol specification
34. Johnson DB, Maltz DA (1996) Dynamic source routing in ad hoc wireless networks. Mob Comput 20:153–181

P. Madadi got his Ph.D from the Electrical and Computer Engineering department of The University of Texas at Austin. He worked under the supervision of Prof. Francois Baccelli and Prof. Gustavo de Veciana. He received his bachelor degree from Indian Institute of Technology, Hyderabad in 2014. His primary research interest is in stochastic modeling of wireless networks. His thesis was focused on analyzing the performance of a mobile user moving in a Poisson cellular network. He is currently working as a senior research engineer at Samsung Research America.

F. Baccelli is Simons Math+X Chair in Mathematics and ECE at UT Austin. His research directions are at the interface between Applied Mathematics (probability theory, stochastic geometry, dynamical systems) and Communications (network science, information theory, wireless networks). He is co-author of research monographs on point processes and queues (with P. Brémaud); max plus algebras and network dynamics (with G. Cohen, G. Olsder and J.P. Quadrat); stationary queuing networks (with P. Brémaud); stochastic geometry and wireless networks (with B. Blaszczyszyn). Before joining UT Austin, he held positions in France, at INRIA, Ecole Normale Supérieure and Ecole Polytechnique. He received the France Télécom Prize of the French Academy of Sciences in 2002 and the ACM Sigmetrics Achievement Award in 2014. He is a co-recipient of the 2014 Stephen O. Rice Prize and of the Leonard G. Abraham Prize Awards of the IEEE Communications Theory Society. He is a member of the French Academy of Sciences and part time researcher at INRIA.

G. de Veciana (S'88-M'94-SM'01-F'09) received his B.S., M.S, and Ph.D. in electrical engineering from the University of California at Berkeley in 1987, 1990, and 1993 respectively, and joined the Department of Electrical and Computer Engineering where he is currently a Cullen Trust Professor of Engineering. He served as the Director and Associate Director of the Wireless Networking and Communications Group (WNCG) at the University of Texas at Austin, from 2003-2007. His research focuses on the analysis and design of communication and computing networks; data-driven decision-making in man-machine systems, and applied probability and queuing theory. Dr. de Veciana served as editor and is currently serving as editor-at-large for the IEEE/ACM Transactions on Networking. He was the recipient of a National Science Foundation CAREER Award 1996 and a co-recipient of five best paper awards including: IEEE William McCalla Best ICCAD Paper Award for 2000, Best Paper in ACM TODAES Jan 2002-2004, Best Paper in ITC 2010, Best Paper in ACM MSWIM 2010, and Best Paper IEEE INFOCOM 2014. In 2009 he was designated IEEE Fellow for his contributions to the analysis and design of communication networks. He currently serves on the board of trustees of IMDEA Networks Madrid