IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received September 3, 2018, accepted October 3, 2018, date of publication October 11, 2018, date of current version November 9, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2875469

Evaluating Performance of RAT Selection Algorithms for 5G Hetnets

DUONG D. NGUYEN[1], HUNG X. NGUYEN[2], (Member, IEEE), AND LANGFORD B. WHITE[1], (Senior Member, IEEE)
[1]School of Electrical and Electronic Engineering, The University of Adelaide, Adelaide, SA 5005, Australia
[2]Teletraffic Research Centre, The University of Adelaide, Adelaide, SA 5005, Australia

Corresponding author: Duong D. Nguyen (duong.nguyen@adelaide.edu.au)

ABSTRACT Next generation 5G cellular networks will consist of multiple technologies for devices to access the network at the edge. One of the keys to 5G is, therefore, the ability of devices to intelligently select its radio access technology (RAT). There have been several proposals for RAT selection in the last few years. Understanding the performance and limitation of these RAT selection solutions is important for their deployment in the future 5G heterogeneous networks. In this paper, we provide a taxonomy and comparative performance analysis of recent RAT selection algorithms, including the different network models that were used to evaluate these works. We combine these different network models to build a benchmark for evaluating the RAT selection algorithms in a 5G environment. We implement the representative algorithms of different approaches and cross compare them in our benchmark. From the experiments conducted, we illustrate how the different network parameters, such as the number of base stations visible to a user and the available link bandwidths, could impact the performance of these algorithms.

INDEX TERMS 5G heterogeneous networks, RAT selection, network models, performance evaluation.

I. INTRODUCTION

The fifth generation of cellular communication (5G) is rapidly gaining momentum worldwide with commercial deployments scheduled for 2020. 5G is expected to offer a variety of novel technologies that can coexist with existing technologies such as 3G and 4G to support diverse requirements of the various applications and services in the future. Heterogeneous networks (HetNets) are built upon multiple wireless access technologies which may include UMTS, GSM, LTE, WiMAX, WiFi, femto as examples [1]. In HetNets, mobile devices with multiple radio access technologies (RATs) are able to choose among the available base stations (BSs). Deciding upon the kind of RAT, and which BS mobile users should connect to is known as the RAT selection problem. This is a topic of considerable ongoing work within the LTE-WLAN interworking framework of the Third Generation Partnership Project (3GPP) [2], and in 5G research [3]–[5].

There is now an extensive body of research on RAT selection solutions in HetNets [6]–[38]. These solutions cover a wide ranges of solution paradigms from centralized to distributed, from one-shot to iterative game theoretic. Most of these works, however, concentrate mainly on developing novel RAT selection algorithms and testing them on specific network topologies or traces. Despite a number of recent surveys of RAT selection techniques [5], thorough comparative performance evaluation of these algorithms under different network settings have not been explored in the literature.

We extend our preliminary study in [38] to provide in this paper a benchmark for studying impact of various network models on the performances of RAT selection algorithms. We mainly focus on evaluating the state-of-the-art algorithms under diverse and realistic network models to understand their strengths and limitation. Our benchmark covers a wide range of network models from throughput, connectivity between users and BSs, and BS deployment. Using this benchmark, we evaluate and cross-compare the performance of the RAT selection algorithms. Our studies illustrate significant performance differences for all algorithms as the various model parameters such as the number of BSs, the number of users and the probability that a link exists between a user and a BS are varied. Significantly, our studies show that the expected number of BSs per user has the most impact on the performance of RAT selection algorithms. Our study indicates that RAT selection algorithms should be evaluated using a range of different network model parameters, perhaps most importantly, the number of BSs available to a user, in order to fully understand their limitations.

Our key contributions are:

1) *A taxonomy of existing RAT selection algorithms:* We firstly undertake a brief survey of existing RAT selection algorithms and evaluation platforms in the literature. Based on their attributes, we classify the algorithms into centralized, distributed and hybrid based approaches. We select and implement representative algorithms from each such group in order to evaluate their performance using multiple metrics, including overall efficiency, system fairness, and convergence behavior.
2) *A unified benchmark for RAT selection algorithms:* We explore different network models used in the evaluation of existing works in order to propose a unified benchmark for performance evaluation of various RAT selection algorithms under the same computational environment and realistic network settings. In particular, we consider two kinds of network models: (i) a random graph based model which represents scenarios where users are distributed independently in the network, and (ii) a geographical based model that more accurately reflects real world scenarios. Our aim is to provide a simulation benchmark for the systematic comparison of different approaches to the RAT selection problem. As far as we know, such a unified framework has not been proposed before.
3) *A thorough comparative study:* This paper provides the first comprehensive evaluation of the effects of different network topology and bandwidth models on the performance of candidate RAT selection algorithms. Issues addressed include user density (the number of users per BS), link density (the number of BSs that a user sees), and bandwidth distribution (the distribution of link bandwidth between BSs and users). To our best knowledge, no comprehensive performance comparison of existing RAT selection algorithms of this kind has been attempted to date.
4) *Software library for RAT selection:* We implement in Matlab a library of different RAT selection algorithms including the default association mechanism using highest signal strength, a centralized algorithm with local search, a wide range of game theoretic algorithms (regret matching, reinforcement learning, non-cooperative scheme, combined fully distributed payoff and strategy reinforcement learning). We make this library publicly available for reference and reuse. The datasets and the codes that we use in this paper, together with their descriptions, are available to access at https://github.com/ndduong1986/RAT-Benchmark.

The rest of this paper is organized as follows. Section II provides a thorough survey of current RAT selection techniques and evaluation platforms. In Section III, we present a unified benchmark for performance evaluation of RAT selection algorithms. The comparative studies and discussions are presented in Section IV. Section V concludes this paper.

II. RAT SELECTION ALGORITHMS AND MODELS

A. RAT SELECTION ALGORITHMS

RAT selection algorithms can be divided into: (i) centralized (network controlled), (ii) distributed (user controlled), or (iii) hybrid (user controlled with network assistance) solutions. We present the most recent state-of-the-art works on the three different approaches. We use BS to denote any network node that connects directly to end users and offers radio access service such as a base station in LTE network or an access point in WiFi.

1) CENTRALIZED RAT SELECTION APPROACHES

In a centralized approach [6]–[15], all the decisions on which RAT a user connects to are made on the network side. In order to do this, all users need to report their local channel conditions to an authorised network controller. Based on this information, the controller calculates the optimal association of users to RATs with respect to a network objective, and then assigns BS to users. Using this centralized mechanism, service providers can maintain control of network operation to achieve some network related objectives such as network throughput maximization [6]–[8], load balancing optimization [9], [10], user fairness enhancement [11], etc. Centralized approach is gaining popularity due to the emergence of future software-defined wireless networks [13]–[15].

Centralized algorithms have been shown to be superior than distributed solutions in term of overall network throughput [16]. They, however, require collaboration between all wireless BSs and users – exchanging significant communication overheads, especially for ultra-dense network deployment [39]. Furthermore, different network operators pursue different network sharing strategies. Therefore, such close collaboration may not be possible across multiple networks.

2) DISTRIBUTED RAT SELECTION APPROACHES

A distributed approach [17]–[23] can overcome the problem of excessive communication overheads, by implementing the RAT selection algorithms at the user side [5]. Most related distributed solutions are iterative game-based algorithms (for a survey refer to [40]). Distributed game-theoretic techniques can be classed into: partially distributed and fully distributed algorithms. A game-theoretic algorithm is considered to be partially distributed if each player (e.g., user) uses information about the other players in order to update its strategy. While using a fully distributed algorithm, players must be able to make decisions without knowledge of the other players (how many there are, their action and payoffs) [41].

In a partially distributed solution such as [17] and [18], for example, to guarantee convergence, all users are assumed to have complete knowledge of the network, including the payoff functions and the history of selection actions of all users.

TABLE 1. Summary of RAT selection algorithms under consideration.

Algorithm	LSH	HSS	RM	RSG	ERL	CODIPAS
Learning-based	Centralized	Fully distributed	Partially distributed	Hybrid	Hybrid	Fully distributed
One-shot or Iterative	One-shot	One-shot	Iterative	Iterative	Iterative	Iterative
Infomation requirements	Global	Local	Global	Global	Local	Local
Data exchange among users	Yes	No	Yes	No	No	No
Knowledge of payoff function	Yes	No	Yes	Yes	No	No
External feedback from BSs	Yes	No	No	Yes	Yes	No
Convergence equilibra	No equilibrium	No equilibrium	CE	NE	CE	NE

From this information, users are able to determine their throughputs given other users' choices. The guaranteed convergence therefore comes at the cost of increased complexity, signaling and communication load.

In contrast, a fully distributed solution such as [19]–[22], for example, does not require the users to have such full knowledge. Each user learns about the RAT selection "game" by observing only its own achieved payoffs. Despite this very attractive property, the conventional fully distributed algorithms in [19]–[22], however, suffer from the problems such as slow convergence, and convergence to sub-optimal equilibrium points, due to the lack of knowledge on global network traffic [23].

3) HYBRID RAT SELECTION APPROACHES

In hybrid approaches [24]–[34], users select their RAT depending on their individual observations as well as external information provided by the network. Several works such as [24]–[28] propose network-assisted schemes where some global knowledge of network is broadcast to every user in the network. Each user then uses these parameters to select the best BS that satisfies its utility requirements. These works however still require significant amounts of additional information exchange between the users and the BSs.

To further reduce the signaling overhead by the broadcast technique, the works in [32]–[34] develop low-overhead distributed algorithms in which each BS shares limited feedback information only to its serving users to assist them in making RAT decision. The feedback sent to the users is related only to the local information of each BS such as the number of connecting users [32], [33], the achievable throughput offered by the BS [33], the BS traffic load [32] or the channel state condition between user and BS [34]. This approach reduces significantly the overheads in the network.

In these hybrid approaches, although the BS may provide some useful information, this knowledge is not guaranteed to be perfect or reflect the global condition of the network. Therefore, users will need to keep switching among the available BSs to discover how it would associate with the BSs to meet its objective. This leads to a high number of exploration times and results in a low per-user throughput.

B. ALGORITHMS UNDER CONSIDERATION

In the following, we discuss and compare the fundamental properties of the six representative algorithms of the classes reviewed above. These algorithms have been considered due to their general approaches and the simplicity of their implementation, as the focus of this research is mainly on the development of a simulation benchmark for comparing main approaches to the RAT selection problem. We limit our discussion to two dominant RATs: WiFi and Cellular. The interworking of these two well-known technologies has been gaining an increasing attention for inter RAT offloading solution in heterogeneous networks, due to their fast deployment and scalable capabilities [4]. Other RATs could be considered within our framework but are beyond the scope of this study. We particularly focus on information input and the types of data exchanged between the users and the BSs. We summarize this discussion in Table 1.

1) ONE-SHOT ALGORITHMS

- **Highest Signal Strength (HSS):** This is a fully distributed approach and is the current default user association mechanism in the 802.11 standard. Users have no information about the global network state. Based on their radio conditions, they randomly select a BS among the highest received signal strengths. To select between WiFi and Cellular RATs, we assume a user randomly belongs to one of the two groups of users. That is either prefer WiFi network or prefer cellular network to mobile access with equal probabilities.
- **Local Search Heuristic (LSH) in [12]:** This algorithm is based on a centralized approach, in which a controller searches for all possible associations between users and BSs. It assigns users to either WiFi or cellular BSs in a way to maximize the sum of logs of the user's throughputs instead of the total users' individual throughput. This optimization method has been shown to significantly improve the overall network throughput while maintains the good fairness of user throughputs.

2) ITERATIVE ALGORITHMS

- **Regret Matching (RM) in [17]:** The motivation behind this partially distributed scheme is to adjust the probabilities over the user's actions to be proportional to the "regrets" for not having chosen alternative actions. In RM, users are assumed to have global information about the network including the history of which BSs have been selected by other users. Each user can compute the regrets (the changes in average payoff) that it would have if choosing other BSs instead of its current selection. To select their RATs, users apply the RM procedure [42] that assures no regret in the long run. This algorithm converges to the set of correlated equilibrium (CE), an optimality concept of game theory that models possible correlation between players, either implicit or explicit. This is in contrast to the usual strategic equilibrium of Nash, where all players act independently [42].

- **RAT Selection Games (RSG) in [26]:** This is a hybrid approach, where a centralized algorithm is used to determine the global network traffic including the number of concurrent users on each BS, and their physical (PHY) data rates. Each BS then broadcasts these parameters to all users in its coverage area, including those that are not currently connected to that BS. Each user can then estimate its expected throughput if it selected another BS. At each time step, each user selects a BS that provides the highest per-user throughput. This algorithm converges to a Nash equilibrium (NE) [26].
- **Enhanced Reinforcement Learning (ERL) in [33]:** This hybrid scheme allows users to estimate their payoffs more accurately using network-assisted feedback. Each BS shares with its serving users, the number of its concurrent users and the long-term achievable throughput (computed at the BS) that a user could receive. From this feedback, and its own observations, each user can estimate its obtainable throughput from all other target BSs and then compute network measured regrets. These quantities indicate how much gain (or loss) in average payoff a user would experience if leaving the currently associated BS. Users then select their RATs by applying the procedure described in [33] (which is based on the regret principle [43]). This algorithm also guarantees convergence to the set of CE almost surely.
- **Combined Fully Distributed Payoff and Strategy Reinforcement Learning (CODIPAS) in [20]:** This is a fully distributed solution where users do not need to exchange their data to other users or BSs. Users adapt their RAT selection decisions only based on their own observation of the payoffs received from past experiences. At each time step, using only this local information, a user selects the best available BS to maximize its payoff. This algorithm guarantees convergence to a NE.

III. A BENCHMARK FOR RAT SELECTION EVALUATION

A. OVERVIEW OF CURRENT EVALUATION PLATFORMS

1) NETWORK TOPOLOGY

There are a large number of network topologies that have been used for wireless network simulations, including both static and dynamic models, e.g. Kauffmann *et. al.* [22]. In a static topology, users in the network are assumed to be static and can only communicate to a fixed set of BSs. This kind of topology is easy to deploy but does not accurately reflect actual network environments. Dynamic topology models represent a more realistic scenario, where users can join or leave the network at any time, but increase the complexity of the simulation model.

Wang *et. al.* [28] evaluate their algorithm in a randomly deployed network, where both BSs and users are distributed according to a homogeneous Poisson Point Process (PPP) in a geographic region. In contrast, Ge *et. al.* [39] use a more complex heterogeneous topology, where BSs are sampled according to a non-homogeneous PPP and therefore results in region with a very high density of BSs. In [6], however, the BS density as well as user density are varied in order to study the impact of these parameters on algorithm performance.

To model the network under different deployment strategies in cellular network, Du *et. al.* [23] use three representative topologies scenarios including a chain-topology (treated as the roadside cellular network BS), a nestification-topology (represented the multimode small cells deployment) and an overlapping-topology (reflected the conventional scenario of partially overlapping cells) to illustrate the applicability of their solution in many complex scenarios. Some other works such as [6], [24], [26], and [28] validate their solutions in real-word networks by using the collected residential data traces via driven experiments.

Typically, most of the existing works evaluate their algorithms on a selected network topology, often with a small number of BSs and full connectivity between users and BSs. These simple models may not reflect the realistic scenarios of future 5G ultra-dense heterogeneous network [39].

2) BANDWIDTH ALLOCATION

Bandwidth allocation is a primary factor that significantly affects performance of wireless network. However, most of the prior works rely on simplifications such as uniform throughput among all clients and consider only a single class of throughput model. For example, the works in [18], [22], and [23] assume that all users connecting to the same BS are allocated with an equal amount of bandwidth. This assumption is suitable to model the throughput-fair access technologies in WLAN environment. Other works apply to a single class of RATs such as WiFi network in [6] and [17] or cellular network in [27] and [28]. Only a number of previous works [24]–[26], [32], [33], [38] look at HetNets scenario where different RATs use different bandwidth allocation techniques. Those solutions that work with multiple RATs are more attractive due to the recent development of HetNets.

Recently, the throughput-fair model and the proportional-fair throughput model in [24]–[26] as well as the service differentiation based throughput model in [23] and [27] are becoming popular. These works however ignore the fluctuating nature of the wireless channel by assuming that the users know the long-term average throughput that a user experiences on a wireless network. Unfortunately, in practice, for distributed or hybrid solutions, each user only knows its sampled throughput (instantaneous value), from which it infers the mean value. Inference from a limited number of samples always contains statistical errors. It is therefore important to take into account the statistical errors in evaluating RAT selection algorithms.

In the rest of this section, we propose a unified simulation model to evaluate and compare the performances of various RAT selection algorithms under the same network environment. With our model, one could also investigate the effect of varying various network parameters on the algorithm performance to fully understand its limitations.

B. NETWORK TOPOLOGY

Consider a wireless network consisting of M BSs and N users. We address two particular classes of such networks: In the first model, associations between users and BSs are based on generic random graph. This model resembles the popular random Poisson point process for users distribution in wireless networks. In the second case, we consider correlated association models based on geographical distance, which better reflect real-word topology deployments.

1) RANDOM GRAPH BASED MODEL

Random graph is a popular mathematical tool to model link connectivity, and to study the scaling capacity of wireless networks [44]. Under a random graph model [44], users are assumed to be located within the coverage range of each BS (hence can potentially connect to that BS) independently of each other with a fixed probability. Random topologies are generated by assigning a probability p that a link (a connection from a user to a BS) is available for a certain user independently among all pairs (user, BS). We define the *link density* as the expected number of BSs (pM), that a user sees. Fig. 1 illustrates the generic random graph scenario.

2) GEOGRAPHICAL BASED MODEL

In this model, the connectivity and bandwidth between BSs and users are determined by their relative geographical distances. We adopt the same network models used in [6] and [28]. These models reflect the real world distributions of BSs and users. We consider a densely deployed networks, where a large number of small cells (e.g., Pico/Femto/WiFi BSs) are located within the coverage area of one macro BS in a narrow area [39]. We divide the given geographic area into smaller, non-overlapping square-shaped areas and randomly placed a BS within the borders of each small area. We then place a uniform random number of users (up to λ, the maximum number of users that a BS serves) for each BS within its area. A user is considered to be a local user to BSs that are located in the same area of its location and to be a non-local users to the rest of the BSs in the network. We assume that each BS can allocate a certain portion of its bandwidth ($0 \leq \alpha \leq 1$), to serve other non-local users ($\alpha = 1$ for the local users).

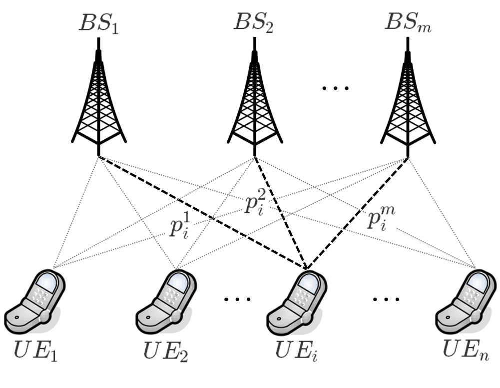


FIGURE 1. The scenario of base stations and users in a random graph based model.

C. BANDWIDTH ALLOCATION

In this paper, we are primarily interested in user downlink throughput and we use the same throughput models as in [24]–[26] for different RATs.

1) THROUGHPUT-FAIR MODEL

Under this model, in the long term, a set of users connected to the same BS receive the same per-user throughput. The throughput for user i when connected to BS k is given by

$$\bar{\omega}_i^k = \left(\sum_{i'=1}^{n^k} \frac{1}{R_{i'}^k} \right)^{-1}, \tag{1}$$

where $R_{i'}^k$ is the PHY rate of user i' on BS k and n^k is the number of users connected to BS k. This model is suitable for throughput-fair access technologies such as WiFi.

2) PROPORTIONAL-FAIR MODEL

Under this model, each user obtains a different throughput which depends on its PHY rate, and the number of other users sharing the same BS. The throughput of user i choosing BS k can be expressed as

$$\bar{\omega}_i^k = \frac{R_i^k}{n^k}. \tag{2}$$

This model is suitable to model time/bandwidth-fair access technologies such as 3G/4G cellular networks.

D. INSTANTANEOUS THROUGHPUT MODEL

The throughputs given in equations (1) and (2) are the mean (e.g., long term average) throughputs, which can be only computed at the network side. In distributed solutions, users only sample their instantaneous throughputs, not the mean values. At any one time, instantaneous throughput observed by the user may vary from the mean. This issue has been considered in [40], where the instantaneous achievable throughput of a user is modeled as a random variable.

In this paper, we propose an instantaneous throughput model that can be efficiently implemented for computer simulations. We assume that the throughput obtained by a user follows a Gaussian distribution in which the mean is equal to the throughput computed by the network and the standard deviation is equal to the product of the noise e and the mean throughput $\bar{\omega}$. Thus, instantaneous throughput of a user i choosing a BS k is a Gaussian random variable:

$$\omega_i^k \sim \mathcal{N}(\bar{\omega}_i^k, \sigma_i^2),$$

where $\sigma_i = e \times \bar{\omega}_i^k$ and $e \in (0, 1)$. This instantaneous throughput model incorporates more practical considerations of real-world networks for RAT selection.

IV. COMPARATIVE STUDIES

We now describe comparative studies of the six algorithms in Section II-B under different network models in Section III-B. We first use synthetic data to simulate a HetNet environment where users are located in the coverage of two different RATs: WiFi and LTE. For the sake of simplicity, it is assumed that one half of the BSs are LTE base stations, while the remaining are WiFi access points. Radio conditions between each user and BS are classified into three levels: good, normal or bad.

Within the scope of this work, we also assume stationary channel conditions over the timescale of the algorithm (i.e., all users remain static during the algorithm's iterations). Considering the impact of user mobility on the performance of different RAT selection algorithms is valid but will be addressed further in our future work. Thus, the PHY rate from a user to a BS is assumed to be unchanged over time in this paper. For each pair of BS k and user i, the good/normal/bad PHY rate to R_i^k is modelled with equal probabilities of 1/3. Table 2 lists the PHY rates when connected alone to these BSs. This assumption will be relaxed by using real network data in Section IV-B.

TABLE 2. PHY rates in WiFi and LTE BSs.

Base station	WiFi	LTE
Good radio condition	48 Mbps	16.6 Mbps
Normal radio condition	24 Mbps	12.2 Mbps
Bad radio condition	9 Mbps	7.4 Mbps

The mean achievable throughput $\bar{\omega}$ a user obtains depends on the other users sharing the same BS, and is given in the equations (1) and (2). When connecting to a BS, a user observes instantaneous throughput ω modelled as a Gaussian random number distributed according to $\mathcal{N}(\mu, \sigma^2)$ with mean $\mu = \bar{\omega}$ and (proportional) variance $\sigma^2 = e^2\,\bar{\omega}^2$, where we assume the proportional noise factor is $e = 0.3$. Results were obtained by averaging over 10 independent trials.

In order to compare the algorithms in term of efficiency and fairness, we consider the following metrics:

- System utility: This is defined as the sum of all users' average throughputs. Higher utilities thus represent greater benefits for both operators (higher offered bandwidth) and users (better per-user throughput).
- Jain's fairness index, which is derived as

$$J = \frac{(\sum_{i=1}^{N} x_i)^2}{N \times \sum_{i=1}^{N} x_i^2}, \quad (3)$$

where x_i is the average throughput of user i and N is the number of users. Notes that J reaches the largest value 1 indicating the best fairness of the system, which guaranteeing the same throughput among the users.

To compare the iterative algorithms in terms of convergence performance, we consider the following metrics:

- Total overheads (bits): This is the amount of data exchanged between users and BSs. A lower overhead is thus preferable.
- Convergence time (iterations): This is the number of iterations required to reach convergence. Fast convergence is desirable because in practical situations, the wireless channel conditions can change quickly.
- Per-user switchings: This is defined to be the maximum number of switchings between RATs required by all users to achieve convergence. A small number of switchings is desirable because each switch occurs an overhead.

A. RANDOM GRAPH BASED MODEL

We first report our results for the random graph case. We vary the link availability probability p between zero and one, and measure the performance of RAT selection algorithms in system fairness and system utility. Figure 2 shows the effect of link density on the performance of the six algorithms described in Section II, for two different numbers of BSs.

1) IMPACT OF LINK DENSITY ON SYSTEM FAIRNESS

Our first observation is that all the iterative algorithms are robust and obtain very good fairness performance compared to that of the one-shot algorithms. This is especially the case when the link density is large ($pM > 4$). Among iterative algorithms, RM, which requires global network information, achieves the best performance. Both regret-based algorithms (RM and ERL) achieve better fairness than the others. This may be explained by noting that both of these algorithms are designed to reach efficient CE points [42], [43], rather than converging to arbitrary NE solutions such as in RSG and CODIPAS. LSH performs worst of the iterative algorithms in term of fairness, with a maximum of 0.85 for a low link density of 6, since it aims to maximize network throughput rather than fairness.

We explain this observation by the following proposition.

Proposition 1: For any constant $\epsilon > 0$, under the random BS and user association model, any user can connect to at least one BS with probability $1 - \epsilon$ if

$$pM \geq M\left(1 - \sqrt[M]{\epsilon}\right),$$

where M is the number of BSs.

Proof: Let random variable X_i denotes the number of BSs that user i sees. Thus, X_i is considered as a binomial random variable with parameter (M, p). The probability for X_i to be l is given by

$$Pr\left[X_i = l\right] = \binom{M}{l} p^l \left(1 - p\right)^{M-l}$$

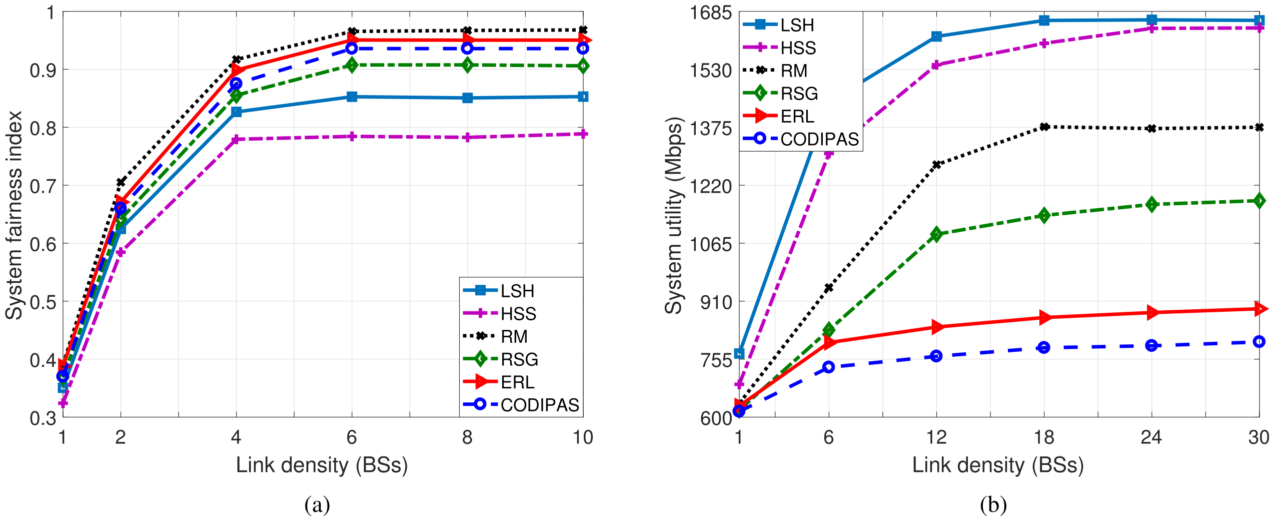


FIGURE 2. **Impact of link density on algorithm performance for the scenario using generic random graph model. (a) 150 users and 10 BSs. (b) 150 users and 30 BSs.**

The probability that user i can see at least one BS can be calculated as

$$Pr\,[X_i \geq 1] = 1 - Pr\,[X_i = 0] = 1 - (1-p)^M \qquad (4)$$

In order to achieve this with high probability, we want

$$Pr[X_i \geq 1] \geq 1 - \epsilon \qquad (5)$$

for all user i. Where ϵ is a pre-determined reliability threshold. For example, to guarantee a 99% confidence interval, we set $\epsilon = 1 - 0.99 = 0.01$. Thus, from (4) and (5), we have

$$1 - (1-p)^M \geq 1 - \epsilon \;\Leftrightarrow\; pM \geq M\left(1 - \sqrt[M]{\epsilon}\right)$$

This completes the proof. ∎

With $M = 10$ and let $\epsilon = 0.01$, we obtain $pM \geq 3.69$ from Proposition 1. We can see that the analytical result match the simulation result reasonable well.

The above formulation means, under probability condition, a user can associate with at least one BS when its link density is higher than a certain threshold value. Accordingly, a distributed iterative algorithm, which aims at maintaining maximum fairness among users, can be used to obtain a high system fairness index at an equilibrium point.

This observation has yielded a primary insight about the impact of link density on the fairness performance of iterative algorithms. That is an iterative game algorithm can achieve very good fairness performance when the link density is large enough (in such dense deployed networks). However, under this scenario, the increase in link density does not help to bring much higher performance in system fairness and therefore can result in wasting network resources.

2) IMPACT OF LINK DENSITY ON SYSTEM UTILITY

In term of utility, the one-shot algorithms achieve much better overall utility than the iterative algorithms. Interestingly, when the link density is 18 in our simulation, the system utility reaches its highest value. Thus, even with higher link density, the centralized LSH algorithm could not bring better system utility. Also, when the link density is large enough ($pM > 24$), the distributed HSS algorithm can achieve similar performance as the centralized one. Thus, in such a densely deployed network, we do not even need a centralized solution in order to maximize overall network throughput.

Among iterative game algorithms, RM that uses global information of the network also achieves highest performance in utility. ERL, which uses assisted feedback only from the local BS, performs worse than RSG, which uses network-assisted information from all BSs, when increasing the link density of the network. CODIPAS, which requires the least amount of network information, has the poorest utility. Again, when the link density reaches 18, the game algorithms could not improve much performance in utility metric.

In any network deployment scenarios, it is important to have mechanisms for associating users to BS so that the available network resources is efficiently used. From the perspective of a single user, increasing network density is always beneficial for increasing individual data rate. However, this might not be optimal from a network-wide viewpoint. In the following, we explore the answer to the question what is the condition for maximizing network throughput. We show in Proposition 2 that even a centralized algorithm (which has complete information regarding the network) could not bring better network throughput when the link density reaches a certain value.

Proposition 2: For any constant $\epsilon > 0$, under the random BS and user association model, total throughput is maximized if

$$pM \geq \frac{M}{\beta}\left(1 - \sqrt[M]{\epsilon}\right),$$

where M is the number of BSs, β is a probability that a user obtains a good radio condition to a BS.

Proof: For simplicity, we assume that a user can obtain a good radio condition to a BS with a fixed probability of β. It is obvious that the network can obtain the maximize throughput if every user can see at least one BS that offers the highest PHY rate (meaning that every user can potentially connect to at least one BS with good radio condition). The probability that a link with good radio condition is available for a certain user is $p\beta$. Let random variable Y_i denotes the number of BSs with good radio conditions that user i can sees. According to binomial distribution,

$$Pr\left[Y_i = l\right] = \binom{M}{l} (p\beta)^l \left(1 - p\beta\right)^{M-l}$$

The probability that user i can see at least one BS with good radio condition can be calculated as

$$Pr\left[Y_i \geq 1\right] = 1 - Pr\left[Y_i = 0\right] = 1 - \left(1 - p\beta\right)^M \tag{6}$$

Similarly, to achieve this with high probability, we want

$$Pr[Y_i \geq 1] \geq 1 - \epsilon \tag{7}$$

for all user i. Thus, from (6) and (7), we have

$$1 - \left(1 - p\beta\right)^M \geq 1 - \epsilon \;\Leftrightarrow\; pM \geq \frac{M}{\beta}\left(1 - \sqrt[M]{\epsilon}\right)$$

When this condition is satisfied, a solution that maximizes the sum of throughput of all the users can be implemented by using a centralized algorithm, such as LSH. The network then can achieve its maximize throughput and hence higher link density does not necessarily provide higher aggressive throughput. This completes the proof. ∎

Let $\beta = 1/3$ according to the simulation setting, Theorem 2 is satisfied for the condition of $pM \geq 11.07$. This result again matches with what we observe in the simulation.

In summary, we observe a similar trend in performance for all algorithms with varying link densities. When the link density is small, increasing the link density of the network brings significant difference in algorithm performance both in fairness and utility. Once the link density reaches a certain threshold, which is 4 in terms of fairness as in Fig. 2(a) and 18 in terms of utility as in Fig. 2(b) in our simulation, all the algorithms reach their limits. Then, higher link density does not necessarily provide higher performance in either fairness or utility and can thus result in wasted network resources. This suggests that neither the number of BSs, nor the probability that a link exists between a user and a BS, has a significant effect on the performance of RAT selection algorithms. It is the link density that is important.

3) CONVERGENCE PERFORMANCE OF ITERATIVE ALGORITHMS

The probability p that a link is available between a user and a BS significantly affects the convergence rate of iterative game-based algorithms. Fig. 3(a) shows the impact of varying p on the convergence speed of RM. As p increases, the convergence rate is observed to improve rapidly. We can also observe a similar impact of the value of p for other schemes (RSG, ERL and CODIPAS). This confirms an obvious property of iterative game-based algorithms: the more information you have, the better the solution.

We now fix $pM = 4$ and evaluate the convergence performances of the four iterative algorithms. Figure 3(b) compares the amount of data exchange (overheads) between users and BSs for different algorithms. Physical quantities such as SNR, PHY rate and throughput are quantized to 4 bit precision. The details of the calculations of the information exchange for each algorithm are summarized below. Here τ denotes the number of iterations to convergence.

- RM: Each user reports its SNR to the connecting BS and records its payoff and PHY rate ($12 \times A$ bits). Each user also records the PHY rates and the actions taken by other $(A-1)$ users in each iteration ($8A(A-1)$ bits). Thus the total overhead is $(8A^2 + 4A) \times$ convergence time (bits) $\sim O(\tau A^2)$.
- RSG: Each user reports its SNR to the BS and records its payoff ($8 \times A$ bits). Then each user records the PHY rates of all the users from each BS ($4 \times A^2$ bits). The total overheads are $(4A^2 + 8A) \times$ convergence time (bits) $\sim O(\tau A^2)$.
- ERL: Apart from the mean achievable throughput (4 bits), Each user records the number of users sharing the same BS (4 bits) and its mean achievable throughput (4 bits). The total overheads are $8A \times$ convergence time (bits) $\sim O(\tau A)$.
- CODIPAS: Each user receives its payoff directly from its associated BS. The total overheads are just $4A \times$ convergence time (bits) $\sim O(\tau A)$.

As illustrated in Fig. 3(b), ERL is the best algorithm in terms of to minimizing overheads. CODIPAS, although requiring less information to make a decision, needs higher overhead because of its slower convergence speed (larger τ). Both ERL and CODIPAS require an order of magnitude less information exchange than RSG and RM, especially when the number of users is large. The reason is that their complexity is linear whereas the complexity of RSG and RM is quadratic.

Figs. 3(c) and 3(d) show the comparison of the algorithms in terms of convergence time and per-user switchings. It can be seen that ERL achieves the fastest convergence rate among all algorithms. Here ERL even outperforms RM and RSG. This can be explained by the fact that the network feedback in ERL is more accurate than the user observed throughput in RM and RSG. Although RM obtains a smaller number of per-user switchings than the other algorithms, it requires a longer time to converge. RM also exchanges significantly higher overheads, as we explained earlier with respect to Fig. 3(b). CODIPAS performs poorest in both speed and per-user switchings metrics due to its lack of global information about network conditions.

B. GEOGRAPHICAL BASED MODEL

To accurately emulate real-work network deployment, we consider an HetNets environment where WiFi BSs and users are located within the coverage area of one macro LTE

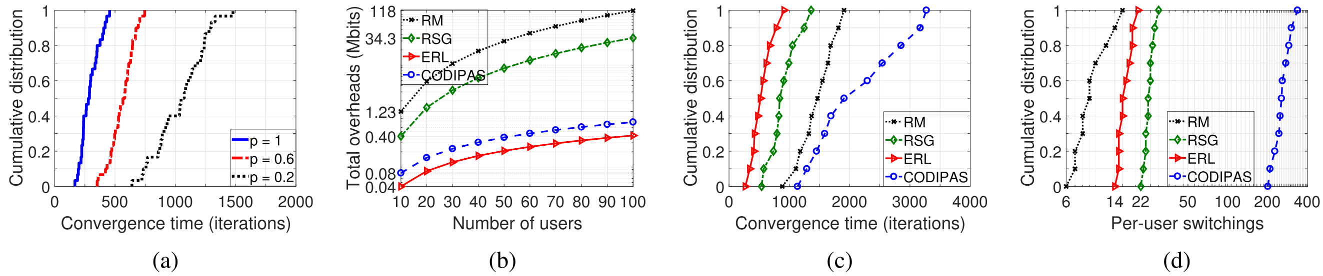


FIGURE 3. (a) Impact of probability of link availability on convergence time of Regret Matching; Convergence performance comparison of iterative algorithms on: (b) Total overheads, (c) Convergence time; (d) Per-user switching.

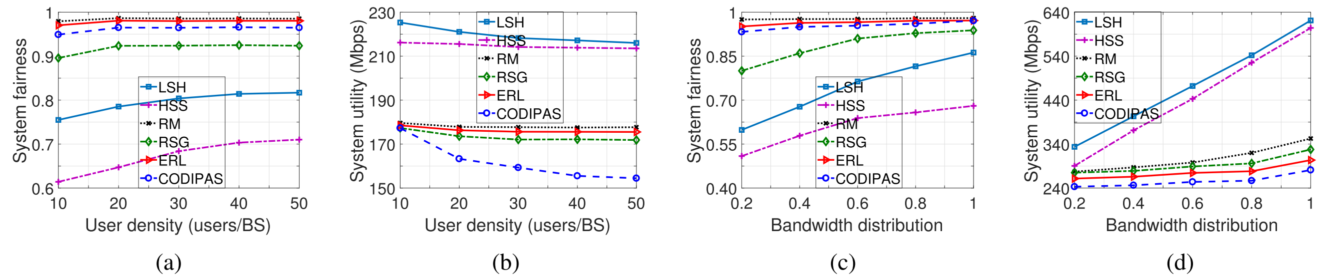


FIGURE 4. (a) Impact of user density on system fairness; (b) Impact of user density on system utility; (c) Impact of bandwidth distribution on system fairness; (d) Impact of bandwidth distribution on system utility.

BS at the center of the network. We use real network data, in particular the measured CQI, from a tier-1 LTE operator to simulate user's PHY rates to the macro LTE BS. In addition to LTE data, we also use the received SNR collected from several WiFi BSs across a university campus, in setting up users' PHY rates to WiFi BSs. These values are then converted to a PHY data rate (which we assume to be constant over time) based on the mapping table of the corresponding technology, and are fed to our simulation. Simulation parameters of the WiFi and the LTE network are set according to [33]. Figs. 4(a) – 4(d) show the impact of the user density (number of user per BS) and bandwidth distribution (portion of bandwidth to serve non-local user) on the performance of the algorithms when using the geographical based model.

1) IMPACT OF USER DENSITY

In this setup, we fix the total number of BS in the network to 5 BSs (composed of 1 LTE BS and 4 WiFi BSs) and enable a share portion of bandwidth $\alpha = 0.3$ on each BS. We vary the user density from 10 to 50. The simulation results are shown in Figs. 4(a) and 4(b).

These figures show that iterative algorithms are quite robust to changes in user density in terms of achieving system fairness, when compared to LHS and HSS. However, the total utility of the network is reduced for all the algorithms as the user density is increased. The reason for this behavior is that rational users running RAT selection algorithms select their BS in order to maximize their own payoffs, which tend to reduce the payoffs of other users connected to the same BS. For example, a user connecting to a BS of low PHY rate could obtain a high payoff when the number of users on that BS is small.

2) IMPACT OF BANDWIDTH DISTRIBUTION

In this setup, we fix the total number of BS in the network to 11 BSs (composed of 1 LTE BS and 10 WiFi BSs) and the user density to 20 users/BS. We vary the bandwidth distribution α on each BS from 0.2 to 1. The simulation results are shown in Figs. 4(c) and 4(d).

These figures show that increasing α improves both utility and fairness. This can be explained by the fact that increasing α is equivalent to increasing the BS density. Thus users have more options to select their preferred BSs that offer the higher PHY rates, which also results in them obtaining better per-user throughputs. Therefore, the overall utility and system fairness improve.

V. CONCLUSION

In this article, we start with a brief review of existing RAT selection algorithms and evaluation platforms. We then investigate the impact of different aspects of network models on the performances of representative algorithms from different approaches via a unified benchmark. Our aims are to compare the performance of various algorithms under the same computational environment and to investigate the effect of various network parameters such as link density, user density and link bandwidth distribution on their performances. The unified evaluation benchmark in this article can serve as a reference for researchers, network developers, or engineers. Both the benchmark and the codes are publicly available for reference purposes.

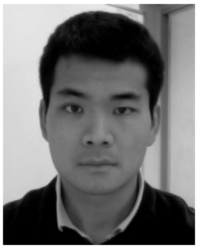
We have studied two particular classes of networks: In the first case, we model random associations between users and base stations in a similar way to the popular random Poisson point process model for the distribution of users in a wireless network. In the second case, we use a correlated association model based on geographical distance that are observed in real world deployments. Simulation results reveal that among all the important network parameters that influence the performance of RAT selection algorithms, the number of base stations that a user can connect to has the most significant impact. This finding provides some guidelines for the proper design of RAT selection algorithms for future 5G.

REFERENCES

[1] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
[2] *3GPP/WLAN RAN Interworking, Release 12*, document TS 37.834, 3rd Generation Partnership Project, Jan. 2014.
[3] A. Ahmed, L. Boulahia, and D. Gaiti, "Enabling vertical handover decisions in heterogeneous wireless networks: A state-of-the-art and a classification," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 2, pp. 776–811, 2nd Quart., 2014.
[4] Y. He, M. Chen, B. Ge, and M. Guizani, "On WiFi offloading in heterogeneous networks: Various incentives and trade-off strategies," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2345–2385, 4th Quart., 2016.
[5] M. Wang, J. Chen, E. Aryafar, and M. Chiang, "A survey of client-controlled HetNets for 5G," *IEEE Access*, vol. 5, pp. 2842–2854, 2016.
[6] O. B. Karimi, J. Liu, and J. Rexford, "Optimal collaborative access point association in wireless networks," in *Proc. INFOCOM IEEE*, May 2014, pp. 1141–1149.
[7] Y. Lin, W. Bao, W. Yu, and B. Liang, "Optimizing user association and spectrum allocation in HetNets: A utility perspective," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1025–1039, Jun. 2015.
[8] D. Bethanabhotla, O. Y. Bursalioglu, H. C. Papadopoulos, and G. Caire, "Optimal user-cell association for massive MIMO wireless networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 1835–1850, Mar. 2016.
[9] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "User association for load balancing in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2706–2716, Jun. 2013.
[10] S. Singh, H. S. Dhillon, and J. G. Andrews, "Downlink rate distribution in multi-RAT heterogeneous networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2013, pp. 5188–5193.
[11] W. Li *et al.*, "AP association for proportional fairness in multirate WLANs," *IEEE/ACM Trans. Netw.*, vol. 22, no. 1, pp. 191–202, Feb. 2014.
[12] M. Amer, A. Busson, and I. G. Lassous, "Association optimization in Wi-Fi networks: Use of an access-based fairness," in *Proc. ACM MSWiM*, Nov. 2016, pp. 119–126.
[13] V. Sagar, R. Chandramouli, and K. P. Subbalakshmi, "Software defined access for HetNets," *IEEE Commun. Mag.*, vol. 54, no. 1, pp. 84–89, Jan. 2016.
[14] T. Han and N. Ansari, "A traffic load balancing framework for software-defined radio access networks powered by hybrid energy sources," *IEEE/ACM Trans. Netw.*, vol. 24, no. 2, no. 2, pp. 1038–1051, Apr. 2016.
[15] S.-N. Yang, S.-W. Ho, Y.-B. Lin, and C.-H. Gan, "A multi-RAT bandwidth aggregation mechanism with software-defined networking," *J. Netw. Comput. Appl.*, vol. 61, pp. 189–198, Feb. 2016.
[16] G. Dandachi, S. Elayoubi, T. Chahed, and N. Chendeb, "Network-centric versus user-centric multihoming strategies in LTE/WiFi networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 4188–4199, May 2016.
[17] L. Chen, "A distributed access point selection algorithm based on no-regret learning for wireless access networks," in *Proc. IEEE 71st Veh. Technol. Conf. (VTC-Spring)*, May 2010, pp. 1–5.
[18] D. Niyato and E. Hossain, "Dynamics of network selection in heterogeneous wireless networks: An evolutionary game approach," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 2008–2017, May 2009.
[19] K. Zhu, D. Niyato, and P. Wang, "Network selection in heterogeneous wireless networks: Evolution with incomplete information," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2010, pp. 1–6.
[20] H. Tembine, "Fully distributed learning for global optima," in *Distributed Strategic Learning for Wireless Engineers*. Boca Raton, FL, USA: CRC Press, 2012, pp. 317–359.
[21] P. Naghavi, S. H. Rastegar, V. Shah-Mansouri, and H. Kebriaei, "Learning RAT selection game in 5G heterogeneous networks," *IEEE Wireless Commun. Lett.*, vol. 5, no. 1, pp. 52–55, Feb. 2016.
[22] B. Kauffmann, F. Baccelli, A. Chaintreau, V. Mhatre, K. Papagiannaki, and C. Diot, "Measurement-based self organization of interfering 802.11 wireless access networks," in *Proc. INFOCOM IEEE*, May 2007, pp. 1451–1459.
[23] Z. Du, Q. Wu, P. Yang, Y. Xu, and Y.-D. Yao, "User-demand-aware wireless network selection: A localized cooperation approach," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4492–4507, Nov. 2014.
[24] E. Aryafar, A. Keshavarz-Haddad, M. Wang, and M. Chiang, "RAT selection games in HetNets," in *Proc. INFOCOM IEEE*, Apr. 2013, pp. 998–1006.
[25] E. Monsef, A. Keshavarz-Haddad, E. Aryafar, J. Saniie, and M. Chiang, "Convergence properties of general network selection games," in *Proc. INFOCOM IEEE*, Apr. 2015, pp. 1445–1453.
[26] A. Keshavarz-Haddad, E. Aryafar, M. Wang, and M. Chiang, "HetNets selection by clients: Convergence, efficiency, and practicality," *IEEE/ACM Trans. Netw.*, vol. 25, no. 1, pp. 406–419, Feb. 2016.
[27] Z. Du, Q. Wu, P. Yang, Y. Xu, J. Wang, and Y.-D. Yao, "Exploiting user demand diversity in heterogeneous wireless networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4142–4155, Aug. 2015.
[28] W. Wang, X. Wu, L. Xie, and S. Lu, "Femto-matching: Efficient traffic offloading in heterogeneous cellular networks," in *Proc. INFOCOM IEEE*, Apr./May 2015, pp. 325–333.
[29] S. Andreev *et al.*, "Intelligent access network selection in converged multi-radio heterogeneous networks," *IEEE Wireless Commun.*, vol. 21, no. 6, pp. 86–96, Dec. 2014.
[30] W. Saad, Z. Han, R. Zheng, M. Debbah, and H. V. Poor, "A college admissions game for uplink user association in wireless small cell networks," in *Proc. INFOCOM IEEE*, Apr./May 2014, pp. 1096–1104.
[31] M. E. Helou, M. Ibrahim, S. Lahoud, K. Khawam, D. Mezher, and B. Cousin, "A network-assisted approach for RAT selection in heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1055–1067, Jun. 2015.
[32] B. H. Jung, N.-O. Song, and D. K. Sung, "A network-assisted user-centric WiFi-offloading model for maximizing per-user throughput in a heterogeneous network," *IEEE Trans. Veh. Technol.*, vol. 63, no. 4, pp. 1940–1945, May 2014.
[33] D. D. Nguyen, H. X. Nguyen, and L. B. White, "Reinforcement learning with network-assisted feedback for heterogeneous RAT selection," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6062–6076, Sep. 2017.
[34] M. Wang, A. Dutta, S. Buccapatnam, and M. Chiang, "Regret-minimizing exploration in HetNets with mmWave," in *Proc. 13th Annu. IEEE Int. Conf. Sens., Commun., Netw. (SECON)*, Jun. 2016, pp. 1–9.
[35] J. Choi, W.-H. Lee, Y.-H. Kim, J.-H. Lee, and S.-C. Kim, "Throughput estimation based distributed base station selection in heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6137–6149, Nov. 2015.
[36] O. Galinina, A. Pyattaev, S. Andreev, M. Dohler, and Y. Koucheryavy, "5G multi-RAT LTE-WiFi ultra-dense small cells: Performance dynamics, architecture, and trends," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1224–1240, Jun. 2015.
[37] S. H. Chae, J.-P. Hong, and W. Choi, "Optimal access in OFDMA multi-RAT cellular networks: Can a single RAT be better?" *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4778–4789, Jul. 2016.
[38] D. D. Nguyen, H. X. Nguyen, and L. B. White, "Performance of adaptive RAT selection algorithms in 5G heterogeneous wireless networks," in *Proc. IEEE 26th Int. Telecommun. Netw. Appl. Conf. (ITNAC)*, Dec. 2016, pp. 70–75.
[39] X. Ge, S. Tu, G. Mao, and C. X. Wang, "5G ultra-dense cellular networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 1, pp. 72–79, Feb. 2016.
[40] Z. Han, D. Niyato, W. Saad, T. Basar, and A. Hjorungnes, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*. Cambridge, U.K.: Cambridge Univ. Press, Oct. 2011.
[41] M. Bennis, M. Simsek, A. Czylwik, W. Saad, S. Valentin, and M. Debbah, "When cellular meets WiFi in wireless small cell networks," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 44–50, Jun. 2013.
[42] S. Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, vol. 68, no. 5, pp. 1127–1150, Sep. 2000.
[43] S. Hart and A. Mas-Colell, "A reinforcement procedure leading to correlated equilibrium," in *Economics Essays*. Berlin, Germany: Springer, 2001, pp. 181–200.
[44] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, Sep. 2009.

DUONG D. NGUYEN received the B.Sc. degree (Hons.) in electronic communication systems from the University of Plymouth in 2008, the M.Sc. degree in mobile and personal communications from King's College London, U.K., in 2009, and the Ph.D. degree in engineering from The University of Adelaide, Australia, in 2018.

He is currently a Post-Doctoral Research Fellow at the School of Electrical and Electronic Engineering, The University of Adelaide. His research interests include signal processing and machine learning techniques for autonomous systems and 5G wireless networks, in particular the application of adaptive reinforcement learning to resource allocation.



HUNG X. NGUYEN (M'11) received the B.Eng. degree in information technology and telecommunications engineering from The University of Adelaide, Australia, and the Ph.D. degree in computer science and telecommunications from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland. He completed the Pre-Doctoral School in communication systems at EPFL.

He joined the Teletraffic Research Centre, The University of Adelaide, in 2012, where he is currently a Senior Research Fellow. His research interests include wireless software-defined networks, cellular technologies, network measurements, tomography, and privacy preserving techniques. He has published over 50 refereed papers on these topics.

LANGFORD B. WHITE (SM'00) received the B.Sc. degree in mathematics and the B.E. (Hons.) and Ph.D. degrees in electrical engineering from The University of Queensland, Australia, in 1984, 1985, and 1989, respectively.

From 1986 to 1999, he was a Research Scientist at the Defence Science and Technology Organisation, Australia, in the areas of radar and communications electronic warfare. Since 1999, he has been a Professor at the School of Electrical and Electronic Engineering, The University of Adelaide, Australia. From 2002 to 2006, he was a fellow with National ICT Australia Ltd. His research interests include the areas of signal processing, control, artificial intelligence, and optimization.

Dr. White was awarded the Australian Telecommunications and Electronics Research Board Medal for Best Young Researcher in 1996.