symmetry

MDPI

Article

Multi-Objective Caching Optimization for Wireless Backhauled Fog Radio Access Network

Alaa Bani-Bakr, MHD Nour Hindia, Kaharudin Dimyati *, Effariza Hanafi * and Tengku Faiz Tengku Mohmed Noor Izam

Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur 50603, Malaysia; alaa.1710@siswa.um.edu.my (A.B.-B.); nourhindia@um.edu.my (M.N.H.); tengkufaiz@um.edu.my (T.F.T.M.N.I.)

* Correspondence: kaharudin@um.edu.my (K.D.); effarizahanafi@um.edu.my (E.H.)

Abstract: Proactive content caching in a fog radio access network (F-RAN) is an efficient technique used to alleviate delivery delay and traffic congestion. However, the symmetric caching of the content is impractical due to the dissimilarity among the contents popularity. Therefore, in this paper, a multi-objective random caching scheme to balance the successful transmission probability (STP) and delay in wireless backhauled F-RAN is proposed. First, stochastic geometry tools are utilized to derive expressions of the association probability, STP, and average delivery delay. Next, the complexity is reduced by considering the asymptotic STP and delay in the high signal-to-noise ratio (SNR) regime. Then, aiming at maximizing the STP or minimizing the delay, the multi-objective cache placement optimization problem is formulated. A novel projected multi-objective cuckoo search algorithm (PMOCSA) is proposed to obtain the Pareto front of the optimal cache placement. The numerical results show that PMOCSA outperforms the original multi-objective cuckoo search algorithm (MOCSA) in terms of convergence to a feasible Pareto front and its rate. It also shows that the proposed multi-objective caching scheme significantly outperforms the well-known benchmark caching schemes by up to 40% higher STP and 85% lower average delay.

Keywords: backhaul; caching; cuckoo search algorithm; delay; fog computing; F-RAN; multi-objective optimization; stochastic geometry; successful transmission probability

Citation: Bani-Bakr, A.; Hindia, M.N.; Dimyati, K.; Hanafi, E.; Tengku Mohmed Noor Izam, T.F. Multi-Objective Caching Optimization for Wireless Backhauled Fog Radio Access Network. *Symmetry* **2021**, *13*, 708. https://doi.org/10.3390/sym13040708

Academic Editors: Simona Halunga and Octavian Fratu

Received: 10 February 2021
Accepted: 12 April 2021
Published: 17 April 2021

1. Introduction

In recent years, cloud radio access networks (C-RANs) have been suffering unprecedented data traffic pressure due to the proliferation of user equipment and the tremendous growth of mobile traffic [1], which is due to the centralized architecture of C-RAN, which leads to many drawbacks including the processing of massive amount of data, high end-to-end delay, and traffic congestion [2]. The fog radio access network (F-RAN) is regarded as a promising solution to alleviate the aforementioned drawbacks owing to its distributed architecture that brings cloud functionalities closer to the end-users at the edge of the network [3]. F-RAN is an extension of C-RAN, in which the fog access points (F-APs) are equipped with limited cache and computing resources. Caching the popular contents at the F-APs can effectively reduce the communication delay and traffic congestion. Therefore, optimizing the content cache placement is of a great importance to improve the performance of F-RAN.

The maximization of the successful transmission probability (STP) [4–8], which is also called success delivery probability or hit probability, and delay minimization [9–18] are the two main addressed cache placement optimization problems in F-RAN. In [4], the caching optimization problem was formulated to maximize the STP in millimeter-wave self-backhauled F-RAN. In [5], the authors proposed an optimal cache placement design for joint and parallel transmission strategies to maximize the STP and the fractional offloaded traffic (FOT). The proactive caching design proposed in [6] utilized the projection gradient

method to maximize the STP in wireless backhauled F-RAN. In [7], a deep Q-learning based content caching scheme in F-RAN was proposed, where the optimization problem was formulated to maximize the hit rate, and the optimal caching matrix was constructed by combing the predicted content popularity and user preference together. In [8], the authors proposed two user preference learning-based edge caching architectures for F-RAN, where the edge caching problem was formulated to maximize the overall cache hit rate. In [9], the authors proposed a hierarchical content caching paradigm for F-RAN, the optimization problem of the proposed paradigm was formulated to alleviate capacity constraints on the fronthaul and to minimize the transmit delay. The centralized and distributed transmission aware cache placement strategies in F-RAN proposed in [10] minimize average download delay of the end-user subject to the cache constraints. In [11], a joint caching and multicast design for wireless fronthaul in F-RAN is proposed, where the optimization problem was formulated to minimize the transmission time for cloud processor. In [12], the authors proposed a socially aware caching scheme for a device-to-device (D2D) enabled F-RAN, where the projective adaptive resonance theory neural network was used to construct the fog community in order to reduce the caching redundancy. Then, the content delivery delay was reduced by using the D2D technology and selecting the most appropriate user equipment to cache data for other users within each access point coverage, where the optimal cache placement at the selected user equipment is obtained using the ant colony optimization algorithm. In [13], the authors proposed a joint radio communication, caching, and computing design for virtual reality delivery in F-RAN, where the joint radio communication, caching, and computing decision optimization problem was formulated to maximize the average tolerant delay. In [14], the problems of content caching, computation offloading, and radio resource allocation in F-RAN were jointly tackled to minimize the average end-to-end delay. In [15], the authors proposed a distributed edge caching strategy in F-RAN, the caching optimization problem of the proposed scheme was formulated as a mean field game aiming at jointly minimizing the request service delay and fronthaul traffic load. The cache update optimization problem to minimize the average transmission delay in downlink F-RAN was addressed in [16]. In [17], the authors proposed a joint proactive caching and power allocation scheme to minimize the delay in a F-RAN, the authors formulated the optimization problem as a mixed-integer nonlinear fractional programming problem, then they utilized the deep Q-learning network to optimize the performance. In [18], the multi-objective caching optimization of the delay and energy efficiency in F-RAN was investigated, where stochastic geometry tools were utilized to formulate the objective functions of the coded caching scheme for multiple transmission strategies. The problem of multi-objective optimization of the STP, FOT, and the delay in cache-enabled F-RAN was addressed in [19], where joint and parallel transmission strategies of coded contents utilizing the cooperative transmission among the F-APs were proposed. The authors in [19] used stochastic geometry tools to formulate the weighted sum multi-objective optimization problem to minimize the delay or to maximize the STP and FOT, then an improved fruit fly optimization algorithm was used to obtain the optimal solution.

The wireless backhauling of the F-APs was not tackled in [5,7–10,12–19]. The impacts of the interference on the hit rate and cache placement were not tackled in [7]. The end user's preferences and quality of service (QoS) were not taken into consideration in [11]. The optimization of the cache placement at the access points was not addressed in [11–13,18,19]. Moreover, the multi-objective optimization of the STP and delay in F-RAN was only investigated in [19], as well as for cooperative coded caching. To the best of the authors knowledge, no prior work has addressed the multi-objective uncoded caching optimization problem of the STP and delay in F-RAN.

Motivated by the aforementioned discussions, this paper proposes a multi-objective optimization to balance the STP and delay in wireless backhauled F-RAN. This work is different from [18,19] not only in terms of considering the uncoded caching, cache placement optimization, and the wireless backhauling of the F-APs, but also the content

dissemination and user association strategies differ. The main contributions of our paper are summarized as follows

1. Closed-form expressions of the probabilities of a F-AP being a direct F-AP or a transit F-AP with respect to the requested content are derived, then the expressions are used to calculate the association probabilities.
2. Using stochastic geometry tools, we derive expressions of the STP and average delay in the general signal-to-interference-plus-noise ratio (SINR) regime. To reduce the complexity, closed-form expressions of the asymptotic multi-objective STP and delay in the high signal-to-noise ratio (SNR) regime are derived.
3. The multi-objective optimization problem is formulated to maximize the STP or minimize the average delay. Then, the asymptotic multi-objective optimization problem in the high SNR is considered to reduce the computational complexity.
4. A novel projected multi-objective cuckoo search algorithm (PMOCSA) is proposed to compute the Pareto front of the optimal cache placement.
5. The numerical results show that the developed PMOCSA outperforms the original multi-objective cuckoo search algorithm (MOCSA). Also, the proposed multi-objective caching scheme is shown to achieve higher performance than the benchmark caching schemes.

The rest of this paper is organized as follows. The system model, including network, caching, and association models, is presented in Section 2. In Section 3, The STP and delay are introduced as a performance metric and analyzed. The problem formulation and optimization are presented in Section 4. The numerical results are delivered and discussed in Section 5. In Section 6, the conclusions are drawn. The key notations used throughout the paper are listed in Table 1.

Table 1. Key notations.

Notation	Description
$\mathcal{M}$	Content library
M	Total number of contents
Φ_F	Point process of the F-APs
Φ_C	Point process of the C-APs
Φ_U	Point process of the users
Φ_m	Point process of the F-APs that cache content m
Φ_{-m}	Point process of the F-APs that do not cache content m
$\Phi_{a,m}$	Point process of the available F-APs with respect to content m
$\Phi_{-a,m}$	Point process of the unavailable F-APs with respect to content m
λ_F	Density of Φ_F
λ_C	Density of Φ_C
λ_U	Density of Φ_U
a_m	Probability of content m being randomly requested by a user
$\boldsymbol{p}$	Caching distribution of the contents
p_m	Probability of content m being cached at each F-AP
b_μ	Probability of the content μ being inactive
Λ_m	Probability of available F-APs with respect to content m
$F_{m,0}$	Direct F-AP with respect to content m
$F_{a,0}$	Transit F-AP with respect to content m
$C_{m,0}$	Nearest C-AP to u_0
C_0	Nearest C-AP to $F_{a,0}$
$\Pr[X_m = F_{m,0}]$	Probability of association with $F_{m,0}$ when content m is requested
$Pr[X_m = F_{a,0}]$	Probability of association with $F_{a,0}$ when content m is requested
$Pr[X_m = C_{m,0}]$	Probability of association with $C_{m,0}$ when content m is requested
$\mathcal{A}_m$	Total probability of association with an access point when content m is requested
$SINR_{m,0}$	SINR at u_0 when it is associated with $F_{m,0}$
$SINR_{a,0}$	SINR at u_0 when it is associated with $F_{a,0}$
$SINR_{C,a}$	SINR at $F_{a,0}$ when u_0 is associated with $F_{a,0}$
$SINR_{C,0}$	SINR at u_0 when it is associated with $C_{m,0}$

Table 1. *Cont.*

Notation	Description
$D_{0,0}$	Distance between $F_{m,0}$ and u_0
$D_{\ell,0}$	Distance between access point ℓ and u_0
$D_{a,0}$	Distance between $F_{a,0}$ and u_0
$D_{C,a}$	Distance between C_0 and $F_{a,0}$
$D_{\ell,a}$	Distance between access point ℓ and $F_{a,0}$
$D_{C,0}$	Distance between $C_{m,0}$ and u_0
$h_{0,0}$	Small-scale channel coefficient between $F_{m,0}$ and u_0
$h_{\ell,0}$	Small-scale channel coefficient between access point ℓ and u_0
$h_{a,0}$	Small-scale channel coefficient between $F_{a,0}$ and u_0
$h_{C,a}$	Small-scale channel coefficient between C_0 and $F_{a,0}$
$h_{\ell,a}$	Small-scale channel coefficient between access point ℓ and $F_{a,0}$
$h_{C,0}$	Small-scale channel coefficient between $C_{m,0}$ and u_0
$q_{m,0}(\boldsymbol{p})$	STP of content m when u_0 is associated with $F_{m,0}$
$q_{m,0,D_{0,0}}(\boldsymbol{p},d)$	$q_{m,0}(\boldsymbol{p})$ conditioned on $D_{0,0}=d$
$q_{C,a,0}(\boldsymbol{p})$	STP of content m when u_0 is associated with $F_{a,0}$
$q_{a,0}(\boldsymbol{p})$	STP of content m over the link $F_{a,0}$ to u_0
$q_{a,0,D_{a,0}}(\boldsymbol{p},d)$	$q_{a,0}(\boldsymbol{p})$ conditioned on $D_{a,0}=d$
$q_{C,a}(\boldsymbol{p})$	STP of content m over the link C_0 to $F_{a,0}$
$q_{C,a,D_{c,a}}(\boldsymbol{p},d)$	$q_{C,a}(\boldsymbol{p})$ conditioned on $D_{C,a}=d$
$q_{C,0}(\boldsymbol{p})$	STP of content m when u_0 is associated with $C_{m,0}$
$q_{C,0,D_{C,0}}(\boldsymbol{p},d)$	$q_{C,0}(\boldsymbol{p})$ conditioned on $D_{C,0}=d$
$q(\boldsymbol{p})$	STP of u_0
$q_\infty(\boldsymbol{p})$	Asymptotic STP of u_0 when $\frac{P}{N_0}\to\infty$
$\tau_{m,0}(\boldsymbol{p})$	Average delay of content m when u_0 is associated with $F_{m,0}$
$\tau_{C,a,0}(\boldsymbol{p})$	Average delay of content m when u_0 is associated with $F_{a,0}$
$\tau_{a,0}(\boldsymbol{p})$	Average delay of content m over the links from $F_{a,0}$ to u_0
$\tau_{C,a}(\boldsymbol{p})$	Average delay of content m over the link from C_0 to $F_{a,0}$
$\tau_{C,0}$	Average delay of content m when u_0 is associated with $C_{m,0}$
$\tau(\boldsymbol{p})$	Average delay of u_0
$\tau_\infty(\boldsymbol{p})$	Asymptotic delay of u_0 when $\frac{P}{N_0}\to\infty$

2. System Model

2.1. Network Model

This paper considers a downlink cache-enabled F-RAN consisting of a tier of cloud access points (C-APs) overlaid with a tier of denser limited storage F-APs. It is assumed that each F-AP is backhauled via a wireless link with the closest C-AP to its location. The F-APs and C-APs are both distributed according to the independent homogeneous Poisson point processes (PPPs) Φ_F and Φ_C of densities λ_F and λ_C, respectively, such that $\lambda_F \gg \lambda_C$. All F-APs and C-APs are assumed to be equipped with a single antenna. The F-APs and C-APs are assumed to transmit at the power of P. The total bandwidth of the F-APs and C-APs are W_F and W_C, respectively. We consider a broadcast transmission scheme, such that each content is disseminated over $1/M_0$ of the total transmission bandwidth of the access point, where M_0 is the total number of contents cached by the access point. The locations of the users are also modeled as an independent homogeneous PPP Φ_U with density λ_U. We assume that each user has a single receive antenna. Let D denote the propagation distance of the transmitted signal. The transmitted signal is assumed to experience a large-scale fading, of which the signal is attenuated by a factor of $D^{-\alpha}$, where α stands for the path loss exponent. The transmitted signal is also assumed to experience a Rayleigh fading, wherein the small-scale fading coefficient h follows an exponential distribution of unit mean (i.e., $|h|^2 \overset{d}{\sim} \exp(1)$).

2.2. Caching Model

Let $\mathcal{M}=\{1,2,...,M\}$ denote the content library, i.e., $\mathcal{M}$ is the set of M contents cached in the network. Due to the limited storage of the F-APs, each F-AP is assumed to cache

a single content in advance. Whereas, all the contents are assumed to be cached in each C-AP. For analytical tractably, the contents are assumed to be of the same size and their popularity distribution among all users is apriori known and identical. Let $\boldsymbol{a} = (a_m)_{m\in\mathcal{M}}$ denote the content popularity distribution, such that $\sum_{m=1}^{M} a_m = 1$ and $a_1 \geq a_2 \geq \cdots \geq a_M$, where $a_m \in (0,1)$ stands for the probability of content m being randomly requested by a user, which is assumed to be characterized by Zipf distribution as follows

$$a_m = \frac{m^{-\gamma}}{\sum_{m\in\mathcal{M}} m^{-\gamma}} \tag{1}$$

where the exponent γ denotes the skew parameter of Zipf distribution.

This paper considers a proactive probabilistic caching strategy of which the caching distribution of the contents is represented by $\boldsymbol{p} = (p_m)_{m\in\mathcal{M}}$, where p_m is the probability of content m being cached at each F-AP, such that it satisfies

$$0 \leq p_m \leq 1, \quad m \in \mathcal{M}, \tag{2}$$

$$\sum_{m\in\mathcal{M}} p_m = 1 \tag{3}$$

2.3. Association Model

Without loss of generality, this paper focuses on a typical user u_0, which is assumed to be located at the origin. Let R denote the discovery range of u_0 and X_m denote the access point that u_0 is associated with when it randomly requests content m. Under the adopted association mechanism illustrated in Figure 1, when the typical user u_0 requests content m, it can be associated with one of the following

1. If there are F-APs caching content m within R, u_0 will be associated with the nearest F-AP caching m $F_{m,0}$ to serve u_0 directly, e.g., user A in Figure 1. Therefore, $F_{m,0}$ is dubbed 'direct F-AP'. Denote $\Pr[X_m = F_{m,0}]$ as the probability of u_0 is being associated with a direct F-AP when it requests content m. As all the F-APs caching content m can serve as a direct F-AP with respect to content m, the point process of the direct F-APs is the thinned PPP $\Phi_m \subseteq \Phi_F$ with density $p_m\lambda_F$, i.e., Φ_m is the point process of the F-APs caching content m. Then, using the null property of PPP, $\Pr[X_m = F_{m,0}]$ can be obtained as in the following lemma

   Lemma 1. *When u_0 requests content m, the probability of u_0 being associated to a direct F-AP within R is given by*

$$\Pr[X_m = F_{m,0}] = 1 - \exp\left(-\pi p_m \lambda_F R^2\right) \tag{4}$$

   Proof. Please refer to Appendix A. □

2. If content m in not cached within R, then u_0 will be associated with the nearest available F-AP $F_{a,0}$ within R to fetch m from the nearest C-AP C_0, e.g., user B in Figure 1. Here, $F_{a,0}$ is dubbed 'transit F-AP' due to the 2-hop transmission. Moreover, the available F-AP is defined as the F-AP that caches inactive content (i.e., a content not requested by users within its association area). Let the random variable $Y_\mu \in \{0,1\}$ denote whether content μ cached by X_μ is being requested by users within its Voronoi cell, such that $Y_\mu = 0$ stands for the event of content μ being not requested, and $Y_\mu = 1$ otherwise. Then, the probability of the content μ being inactive can be obtained using proposition 1 of [20] as

$$b_\mu = Pr\left[Y_\mu = 0\right] = \left(1 + \frac{a_\mu\, \lambda_U}{3.5\, p_\mu\, \lambda_F}\right)^{-3.5} \tag{5}$$

The probability that a F-AP can be an available F-AP Λ_m when content m is requested can be obtained as in the following lemma

Lemma 2. *The probability of available F-APs with respect to content m is*

$$\Lambda_m = \sum_{\mu \in \mathcal{M} \setminus m} p_\mu b_\mu \tag{6}$$

Proof. Please refer to Appendix B. □

To proceed, we denote $\Phi_{a,m} \subseteq \Phi_F$ as the point process of available F-APs with respect to content m. As $\Phi_{a,m}$ is a thinned PPP, its density can be obtained as $\Lambda_m \lambda_F$. Then, the probability of the event that u_0 is associated with a transit F-AP within R when it requests content m can be obtained as in Lemma 3

Lemma 3. *When u_0 requests content m, the probability of u_0 is being associated with a transit F-AP within R is given by*

$$Pr[X_m = F_{a,0}] = \exp\left(-\pi p_m \lambda_F R^2\right)\left(1 - \exp\left(-\pi \Lambda_m \lambda_F R^2\right)\right) \tag{7}$$

Proof. Please refer to Appendix C. □

3. If neither a direct nor a transit F-AP exists within R to be associated with, then u_0 will be associated with the nearest C-AP $C_{m,0}$ within R, e.g., user C in Figure 1. The probability of this event $Pr[X_m = C_{m,0}]$ is given as in the following lemma

Lemma 4. *When u_0 requests content m, the probability of u_0 is being associated with the nearest C-AP within R is given by*

$$Pr[X_m = C_{m,0}] = \exp\left(-\pi (p_m + \Lambda_m) \lambda_F R^2\right)\left(1 - \exp\left(-\pi \lambda_C R^2\right)\right) \tag{8}$$

Proof. Please refer to Appendix D. □

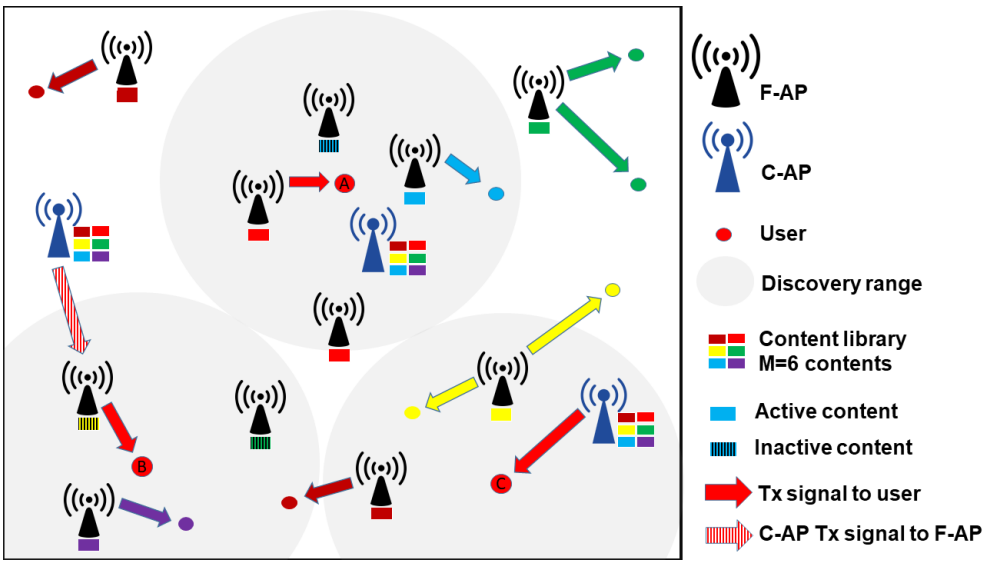

Figure 1. Association model.

3. Performance Analysis

In this section, the STP and delay are analyzed to evaluate the F-RAN performance. The STP represents the probability that a requested content can be successfully transmitted. Whereas, the delay represents the average time of a successful content delivery.

3.1. STP Analysis

When u_0 requesting content m is being associated with the direct F-AP $F_{m,0}$ within R, the requested content m can be successfully received and decoded at a transmission rate ξ, if the channel capacity of u_0 is greater than or equal to ξ. Thus, the STP of content m when u_0 is being associated with $F_{m,0}$ can be expressed as

$$\begin{aligned} q_{m,0}(\boldsymbol{p}) &= \Pr[\mathfrak{R}_{m,0} \geq \xi] \\ &= \Pr[W_F \log_2(1 + SINR_{m,0}) \geq \xi] \end{aligned} \tag{9}$$

where $\mathfrak{R}_{m,0} = W_F \log_2(1 + SINR_{m,0})$ is the channel capacity of link between u_0 and $F_{m,0}$, and $SINR_{m,0}$ is the SINR of u_0 when it is associated with $F_{m,0}$ given by

$$SINR_{m,0} = \frac{D_{0,0}^{-\alpha} |h_{0,0}|^2}{\sum_{\ell \in \Phi_m \setminus F_{m,0}} D_{\ell,0}^{-\alpha} |h_{\ell,0}|^2 + \sum_{\ell \in \Phi_{-m}} D_{\ell,0}^{-\alpha} |h_{\ell,0}|^2 + \sum_{\ell \in \Phi_C} D_{\ell,0}^{-\alpha} |h_{\ell,0}|^2 + \frac{N_0}{P}} \tag{10}$$

where Φ_{-m} denotes the point process of the F-APs do not cache content m, $D_{0,0}$ and $D_{\ell,0}$ are the distance between u_0 and the direct F-AP $F_{m,0}$, and the distance between u_0 and access point ℓ, respectively. $h_{0,0}$ and $h_{\ell,0}$ are the small-scale channel coefficients between u_0 and $F_{m,0}$, and u_0 and access point ℓ, respectively. N_0 denotes the noise power. Next, stochastic geometry tools are utilized to obtain $q_{m,0}(\boldsymbol{p})$ as given in Theorem 1.

Theorem 1. *The STP of content m when u_0 is associated with $F_{m,0}$ is given by*

$$\begin{aligned} q_{m,0}(\boldsymbol{p}) &= \int_0^R \underbrace{\exp\left(-\pi p_m \lambda_F \mathcal{U}(\boldsymbol{p}) d^2 - \frac{N_0}{P}\left(2^{\frac{\xi}{W_F}} - 1\right) d^\alpha\right)}_{= q_{m,0,D_{0,0}}(\boldsymbol{p},d)} \underbrace{2\pi p_m \lambda_F d \exp\left(-\pi p_m \lambda_F d^2\right)}_{= f_{D_{0,0}}(d)} dd \\ &= 2\pi p_m \lambda_F \int_0^R d \exp\left(-\pi p_m \lambda_F \left(1 + \mathcal{U}(\boldsymbol{p})\right) d^2 - \frac{N_0}{P}\left(2^{\frac{\xi}{W_F}} - 1\right) d^\alpha\right) dd \end{aligned} \tag{11}$$

where $q_{m,0,D_{0,0}}$ is the conditional STP (i.e., $q_{m,0}$ conditioned on the distance $D_{0,0} = d$), $f_{D_{0,0}}(d)$ is the probability density function (PDF) of $D_{0,0}$, and

$$\mathcal{U}(\boldsymbol{p}) = \frac{2}{\alpha}\left(2^{\frac{\xi}{W_F}} - 1\right)^{\frac{2}{\alpha}} \left(\beta'\left(\frac{2}{\alpha}, 1 - \frac{2}{\alpha}, 2^{\frac{-\xi}{W_F}}\right) + \frac{\lambda_F - p_m \lambda_F + \lambda_C}{p_m \lambda_F} \beta\left(\frac{2}{\alpha}, 1 - \frac{2}{\alpha}\right)\right) \tag{12}$$

where $\beta'(x, y, z) \triangleq \int_z^1 u^{x-1}(1-u)^{y-1} du$ is the complementary incomplete Beta function and $\beta(x, y) \triangleq \int_0^1 u^{x-1}(1-u)^{y-1} du$ is the Beta function.

Proof. Please refer to Appendix E. □

The STP of content m when u_0 is associated with the transit F-AP $F_{a,0}$ can be formulated due to the 2-hop transmission as

$$\begin{aligned} q_{C,a,0}(\boldsymbol{p}) &= q_{a,0}(\boldsymbol{p}) q_{C,a}(\boldsymbol{p}) \\ &= \Pr\left[\underbrace{W_F \log_2(1 + SINR_{a,0})}_{\triangleq \mathfrak{R}_{a,0}} \geq \xi\right] \Pr\left[\underbrace{\frac{W_C}{M} \log_2(1 + SINR_{C,a})}_{\triangleq \mathfrak{R}_{C,a}} \geq \xi\right] \end{aligned} \tag{13}$$

where $q_{a,0}(\boldsymbol{p})$ and $q_{C,a}(\boldsymbol{p})$ are the STPs of content m over the links $F_{a,0}$ to u_0, and C_0 to $F_{a,0}$, respectively. $\mathfrak{R}_{a,0}$ and $\mathfrak{R}_{C,a}$ are the channel capacities of the links $F_{a,0}$ to u_0, and C_0 to $F_{a,0}$, respectively. $SINR_{a,0}$ is the SINR at u_0 and $SINR_{C,a}$ is the SINR at $F_{a,0}$, which are given by

$$SINR_{a,0} = \frac{D_{a,0}^{-\alpha}\,|h_{a,0}|^2}{\sum_{\ell\in\Phi_{a,m}\setminus F_{a,0}} D_{\ell,0}^{-\alpha}\,|h_{\ell,0}|^2 + \sum_{\ell\in\Phi_{-a,m}} D_{\ell,0}^{-\alpha}\,|h_{\ell,0}|^2 + \sum_{\ell\in\Phi_C} D_{\ell,0}^{-\alpha}\,|h_{\ell,0}|^2 + \frac{N_0}{P}} \tag{14}$$

and

$$SINR_{C,a} = \frac{D_{C,a}^{-\alpha}\,|h_{C,a}|^2}{\sum_{\ell\in\Phi_C\setminus C_0} D_{\ell,a}^{-\alpha}\,|h_{\ell,a}|^2 + \sum_{\ell\in\Phi_F\setminus F_{a,0}} D_{\ell,a}^{-\alpha}\,|h_{\ell,a}|^2 + \frac{N_0}{P}} \tag{15}$$

where $\Phi_{-a,m} \triangleq \Phi_F \setminus \Phi_{a,m}$ with density $(1-\Lambda_m)\lambda_F$ denotes the point process of unavailable F-APs with respect to content m. $D_{a,0}$, $D_{C,a}$, and $D_{\ell,a}$ are the distances between $F_{a,0}$ and u_0, C_0, and access point ℓ, respectively. $h_{a,0}$, $h_{C,a}$, and $h_{\ell,a}$ are the small-scale channel coefficients of the links $F_{a,0}$ to u_0, and C_0 to $F_{a,0}$, and access point ℓ to $F_{a,0}$, respectively. Then, stochastic geometry is utilized to obtain the expression of $q_{C,a,0}(\boldsymbol{p})$ in the following theorem.

Theorem 2. *The STP of content m when u_0 is associated with $F_{a,0}$ can be expressed as*

$$\begin{aligned} q_{C,a,0}(\boldsymbol{p}) &= \underbrace{\int_0^R \underbrace{\exp\left(-\pi\Lambda_m\lambda_F\mathcal{V}(\boldsymbol{p})d^2 - \frac{N_0}{P}\left(2^{\frac{\xi}{W_F}}-1\right)d^\alpha\right)}_{=q_{a,0,D_{a,0}}(\boldsymbol{p},d)} \underbrace{2\pi\Lambda_m\lambda_F d\exp\left(-\pi\Lambda_m\lambda_F d^2\right)}_{=f_{D_{a,0}}(d)} dd}_{=q_{a,0}(\boldsymbol{p})} \\ &\quad \times \underbrace{\int_0^\infty \underbrace{\exp\left(-\pi\lambda_C\mathcal{G}\,d^2 - \frac{N_0}{P}\left(2^{\frac{M\xi}{W_C}}-1\right)d^\alpha\right)}_{=q_{C,a,D_{C,a}}(\boldsymbol{p},d)} \underbrace{2\pi\lambda_C d\exp\left(-\pi\lambda_C d^2\right)}_{=f_{D_{C,a}}(d)} dd}_{=q_{C,a}(\boldsymbol{p})} \\ &= 2\pi\Lambda_m\lambda_F\int_0^R d\exp\left(-\pi\Lambda_m\lambda_F(1+\mathcal{V}(\boldsymbol{p}))\,d^2 - \frac{N_0}{P}\left(2^{\frac{\xi}{W_F}}-1\right)d^\alpha\right)dd \\ &\quad \times 2\pi\lambda_C\int_0^\infty d\exp\left(-\pi\lambda_C(1+\mathcal{G})\,d^2 - \frac{N_0}{P}\left(2^{\frac{M\xi}{W_C}}-1\right)d^\alpha\right)dd \end{aligned} \tag{16}$$

where $q_{a,0,D_{a,0}}(\boldsymbol{p},d)$ and $q_{C,a,D_{c,a}}(\boldsymbol{p},d)$ represent the conditional STPs conditioned on $D_{a,0}=d$ and $D_{C,a}=d$, respectively. $f_{D_{a,0}}(d)$ and $f_{D_{C,a}}(d)$ are the PDF of $D_{a,0}$ and $D_{C,a}$, respectively.

$$\mathcal{V}(\boldsymbol{p}) = \frac{2}{\alpha}\left(2^{\frac{\xi}{W_F}}-1\right)^{\frac{2}{\alpha}}\left(\beta'\left(\frac{2}{\alpha},1-\frac{2}{\alpha},2^{\frac{-\xi}{W_F}}\right) + \frac{\lambda_F-\Lambda_m\lambda_F+\lambda_C}{\Lambda_m\lambda_F}\beta\left(\frac{2}{\alpha},1-\frac{2}{\alpha}\right)\right) \tag{17}$$

and

$$\mathcal{G} = \frac{2}{\alpha}\left(2^{\frac{M\xi}{W_C}}-1\right)^{\frac{2}{\alpha}}\left(\beta'\left(\frac{2}{\alpha},1-\frac{2}{\alpha},2^{\frac{-M\xi}{W_C}}\right) + \frac{\lambda_F}{\lambda_C}\beta\left(\frac{2}{\alpha},1-\frac{2}{\alpha}\right)\right) \tag{18}$$

Proof. Please refer to Appendix F. □

The STP of content m when the requester u_0 is associated with nearest C-AP $C_{m,0}$ within R is given as

$$q_{C,0}(\boldsymbol{p}) = \Pr\Big[\underbrace{\frac{W_C}{M}\log_2(1+SINR_{C,0})}_{\triangleq\mathfrak{R}_{C,0}} \geq \xi\Big] \tag{19}$$

where $\mathfrak{R}_{C,0}$ denotes the channel capacity of the link from $C_{m,0}$ to u_0, and $SINR_{C,0}$ is the SINR at u_0 when it is associated with $C_{m,0}$ which is given as

$$SINR_{C,0} = \frac{D_{C,0}^{-\alpha} |h_{C,0}|^2}{\sum_{\ell \in \Phi_C \setminus C_{m,0}} D_{\ell,0}^{-\alpha} |h_{\ell,0}|^2 + \sum_{\ell \in \Phi_F} D_{\ell,0}^{-\alpha} |h_{\ell,0}|^2 + \frac{N_0}{P}} \tag{20}$$

where $D_{C,0}$ is the distance between u_0 and $C_{m,0}$, and $h_{C,0}$ is the small-scale channel coefficient between u_0 and $C_{m,0}$. Utilizing stochastic geometry, $q_{C,0}$ can be computed as given in Theorem 3.

Theorem 3. *The STP of content m when u_0 is associated with $C_{m,0}$ can be calculated by*

$$q_{C,0}(\boldsymbol{p}) = \int_0^R \underbrace{\exp\left(-\pi\lambda_C \mathcal{G}\, d^2 - \frac{N_0}{P}\left(2^{\frac{M\xi}{W_C}} - 1\right) d^\alpha\right)}_{=q_{C,0,D_{C,0}}(\boldsymbol{p},d)} \underbrace{2\pi\lambda_C\, d \exp\left(-\pi\lambda_C\, d^2\right)}_{=f_{D_{C,0}}(d)} dd$$

$$= 2\pi\lambda_C \int_0^R d \exp\left(-\pi\lambda_C (1+\mathcal{G})\, d^2 - \frac{N_0}{P}\left(2^{\frac{M\xi}{W_C}} - 1\right) d^\alpha\right) dd \tag{21}$$

where $q_{C,0,D_{C,0}}(\boldsymbol{p}, d)$ denotes the conditional STP conditioned on $D_{C,0} = d$, $f_{D_{C,0}}(d)$ is the PDF of $D_{C,0}$, and $\mathcal{G}$ is given by (18).

Proof. Please refer to Appendix G. □

Since there are M different contents and three different association schemes to deliver the requested content, the STP of u_0 can be obtained using total probability theorem as in Theorem 4.

Theorem 4. *The STP of u_0 can be calculated as*

$$q(\boldsymbol{p}) = \sum_{m \in \mathcal{M}} a_m \Big(\Pr[X_m = F_{m,0}]\, q_{m,0}(\boldsymbol{p}) + \Pr[X_m = F_{a,0}]\, q_{C,a,0}(\boldsymbol{p}) + \Pr[X_m = C_{m,0}]\, q_{C,0}(\boldsymbol{p}) \Big) \tag{22}$$

where $\Pr[X_m = F_{m,0}]$, $\Pr[X_m = F_{a,0}]$, $\Pr[X_m = C_{m,0}]$, $q_{m,0}$, $q_{C,a,0}$, and $q_{C,0}$ are given in Lemmas 1, 3 and 4, and Theorems 1–3, respectively.

It is noted that $q(\boldsymbol{p})$ is very complex. Therefore, the complexity is reduced by considering the asymptotic STP in the high SNR regime (i.e., $\frac{P}{N_0} \to \infty$) as given in Corollary 1.

Corollary 1 (Asymptotic STP). *When $\frac{P}{N_0} \to \infty$, we have*

$$
\begin{aligned}
q_{\infty}(\boldsymbol{p}) &\triangleq \lim_{\frac{P}{N_0}\to\infty} q(\boldsymbol{p}) \\
&= \sum_{m\in\mathcal{M}} a_m \Bigg(\Big(\Pr[X_m = F_{m,0}] \times 2\pi p_m \lambda_F \int_0^R d \exp\Big(-\pi p_m \lambda_F \left(1+\mathcal{U}(\boldsymbol{p})\right) d^2\Big) dd \Big) \\
&\qquad + \Bigg(\Pr[X_m = F_{a,0}] \times 2\pi \Lambda_m \lambda_F \int_0^R d \exp\Big(-\pi \Lambda_m \lambda_F (1+\mathcal{V}(\boldsymbol{p}))\, d^2\Big) dd \\
&\qquad\qquad \times 2\pi\lambda_C \int_0^{\infty} d \exp\Big(-\pi\lambda_C (1+\mathcal{G})\, d^2\Big) dd \Bigg) \\
&\qquad + \Big(\Pr[X_m = C_{m,0}] \times 2\pi\lambda_C \int_0^R d \exp\Big(-\pi\lambda_C (1+\mathcal{G})\, d^2\Big) dd \Big) \Bigg) \\
&= \sum_{m\in\mathcal{M}} a_m \Big(\Pr[X_m = F_{m,0}] \frac{1-\exp\left(-\pi p_m \lambda_F \left(1+\mathcal{U}(\boldsymbol{p})\right) R^2\right)}{1+\mathcal{U}(\boldsymbol{p})} \\
&\qquad + \Pr[X_m = F_{a,0}] \frac{1-\exp\left(-\pi \Lambda_m \lambda_F \left(1+\mathcal{V}(\boldsymbol{p})\right) R^2\right)}{(1+\mathcal{G})\left(1+\mathcal{V}(\boldsymbol{p})\right)} \\
&\qquad + \Pr[X_m = C_{m,0}] \frac{1-\exp\left(-\pi\lambda_C \left(1+\mathcal{G}\right) R^2\right)}{1+\mathcal{G}} \Big)
\end{aligned} \tag{23}
$$

where $\mathcal{U}(\boldsymbol{p})$, $\mathcal{V}(\boldsymbol{p})$, and $\mathcal{G}$ are given in (12), (17), and (18), respectively.

Proof. Since the terms containing $\frac{N_0}{P}$ in $q_{m,0}$, $q_{a,0}$, $q_{C,a}$ and $q_{C,0}$ approach zero as $\frac{P}{N_0} \to \infty$, $q_{m,0}$, $q_{a,0}$, $q_{C,a}$, and $q_{C,0}$ reduce to be in the following form $\int_0^{c_1} c_2 d \exp\left(-c_3 d^2\right) \mathrm{d}d$, where c_1, c_2, and c_3 are constants. Then, the solution of the integral can be computed by $\frac{c_2}{2c_3}\left(1-\exp\left(-c_3 c_1^2\right)\right)$ and thus we can prove Corollary 1. □

3.2. Delay Analysis

The considered average delay in this paper is defined as the average required time to successfully receive the requested content by the typical user. Due to the retransmission of the requested content if an outage event occurs in a time slot, the average delay is correlated with the average number of time slots required to successfully receive the requested content, which is a geometric variable. Thus, when the typical user is associated with $F_{m,0}$, the average delay of content m can be expressed as

$$
\tau_{m,0}(\boldsymbol{p}) = T\, E_{D_{m,0}}\left[\frac{1}{q_{m,0,D_{m,0}}(\boldsymbol{p}, d)}\right] \tag{24}
$$

here, T denotes the duration of the time slot, and $E_{D_{m,0}}\left[1/q_{m,0,D_{m,0}}(\boldsymbol{p}, d)\right]$ represents the average number of required time slots to successfully receive content m. Next, $\tau_{m,0}(\boldsymbol{p})$ can be obtained as in the following theorem.

Theorem 5. *The average delay of content m when u_0 is associated with $F_{m,0}$ is given by*

$$
\begin{aligned}
\tau_{m,0}(\boldsymbol{p}) &= T \int_0^R \underbrace{\exp\left(\pi p_m \lambda_F \mathcal{U}(\boldsymbol{p}) d^2 + \frac{N_0}{P}\left(2^{\frac{\xi}{W_F}}-1\right) d^{\alpha}\right)}_{=\frac{1}{q_{m,0,D_{0,0}}(\boldsymbol{p},d)}} \underbrace{2\pi p_m \lambda_F d\, \exp\left(-\pi p_m \lambda_F d^2\right)}_{=f_{D_{0,0}}(d)} dd \\
&= 2\pi p_m \lambda_F T \int_0^R d \exp\left(-\pi p_m \lambda_F \left(1-\mathcal{U}(\boldsymbol{p})\right) d^2 + \frac{N_0}{P}\left(2^{\frac{\xi}{W_F}}-1\right) d^{\alpha}\right) dd
\end{aligned} \tag{25}
$$

Proof. The proof is straightforward, noting that $E_{D_{m,0}}\left[1/q_{m,0,D_{m,0}}(\boldsymbol{p},d)\right] = \int_0^R (1/q_{m,0,D_{0,0}}(\boldsymbol{p},d))\, f_{D_{0,0}}(d)\mathrm{d}d$, where $q_{m,0,D_{m,0}}(\boldsymbol{p},d)$ and $f_{D_{0,0}}(d)$ are given in Theorem 1. □

In the same manner, when u_0 is associated with the transit F-AP, the average delay of the requested content m can be expressed as in Theorem 6.

Theorem 6. *The average delay of content m when u_0 is associated with $F_{a,0}$ can be calculated as*

$$\begin{aligned}\tau_{C,a,0}(\boldsymbol{p}) &= \underbrace{T\int_0^R \underbrace{\exp\left(\pi\Lambda_m\lambda_F\mathcal{V}(\boldsymbol{p})d^2+\frac{N_0}{P}\left(2^{\frac{\xi}{W_F}}-1\right)d^\alpha\right)}_{=\frac{1}{q_{a,0,D_{a,0}}(\boldsymbol{p},d)}}\underbrace{2\pi\Lambda_m\lambda_F d\exp\left(-\pi\Lambda_m\lambda_F d^2\right)}_{=f_{D_{a,0}}(d)}dd}_{=\tau_{a,0}(\boldsymbol{p})}\\ &\quad+\underbrace{T\int_0^\infty \underbrace{\exp\left(\pi\lambda_C\mathcal{G}\,d^2+\frac{N_0}{P}\left(2^{\frac{M\xi}{W_C}}-1\right)d^\alpha\right)}_{=\frac{1}{q_{C,a,D_{C,a}}(\boldsymbol{p},d)}}\underbrace{2\pi\lambda_C d\exp\left(-\pi\lambda_C\,d^2\right)}_{=f_{D_{C,a}}(d)}dd}_{=\tau_{C,a}(\boldsymbol{p})}\\ &= 2\pi\Lambda_m\lambda_F T\int_0^R d\exp\left(-\pi\Lambda_m\lambda_F(1-\mathcal{V}(\boldsymbol{p}))\,d^2+\frac{N_0}{P}\left(2^{\frac{\xi}{W_F}}-1\right)d^\alpha\right)dd\\ &\quad+2\pi\lambda_C T\int_0^\infty d\exp\left(-\pi\lambda_C(1-\mathcal{G})\,d^2+\frac{N_0}{P}\left(2^{\frac{M\xi}{W_C}}-1\right)d^\alpha\right)dd \end{aligned} \tag{26}$$

where $\tau_{C,a}(\boldsymbol{p})$ and $\tau_{a,0}(\boldsymbol{p})$ are the average delays of content m over the links from C_0 to $F_{a,0}$, and $F_{a,0}$ to u_0, respectively.

Proof. Due to the 2-hop transmission, the average delay can be formulated as $\tau_{C,a,0}(\boldsymbol{p}) = \tau_{a,0}(\boldsymbol{p}) + \tau_{C,a}(\boldsymbol{p})$. Then, by taking the expectation of conditional STPs $q_{a,0,D_{a,0}}$ and $q_{C,a,D_{C,a}}$, the average delay can be calculated as in (26), where $q_{a,0,D_{a,0}}$, $q_{C,a,D_{C,a}}$, $f_{D_{a,0}}(d)$, and $f_{D_{C,a}}(d)$ are given in Theorem 2. □

The average delay of the requested content m when u_0 downloads it from the nearest C-AP is given in the following theorem.

Theorem 7. *The average delay of content m when u_0 is associated with $C_{m,0}$ can be expressed as*

$$\begin{aligned}\tau_{C,0} &= T\int_0^R \underbrace{\exp\left(\pi\lambda_C\mathcal{G}\,d^2+\frac{N_0}{P}\left(2^{\frac{M\xi}{W_C}}-1\right)d^\alpha\right)}_{=\frac{1}{q_{C,0,D_{C,0}}(\boldsymbol{p},d)}}\underbrace{2\pi\lambda_C\,d\exp\left(-\pi\lambda_C\,d^2\right)}_{=f_{D_{C,0}}(d)}dd\\ &= 2\pi\lambda_C T\int_0^R d\exp\left(-\pi\lambda_C(1-\mathcal{G})\,d^2+\frac{N_0}{P}\left(2^{\frac{M\xi}{W_C}}-1\right)d^\alpha\right)dd \end{aligned} \tag{27}$$

Proof. The expectation $E_{D_{C,0}}\left[1/q_{C,0,D_{C,0}}(\boldsymbol{p},d)\right]$ can be calculated by $\int_0^R (1/q_{C,0,D_{C,0}}(\boldsymbol{p},d)) \times f_{D_{C,0}}(d)\mathrm{d}d$, where $q_{C,0,D_{C,0}}(\boldsymbol{p},d)$ and $f_{D_{C,0}}(d)$ are provided in Theorem 3. □

Using the total probability theorem, the average delay of the typical user can be calculated as in the following theorem.

Theorem 8. *The average delay of u_0 is given as*

$$\tau(\boldsymbol{p}) = T\sum_{m\in\mathcal{M}}\frac{a_m}{\mathcal{A}_m}\left(\Pr[X_m=F_{m,0}]\,\tau_{m,0}(\boldsymbol{p})+\Pr[X_m=F_{a,0}]\,\tau_{C,a,0}(\boldsymbol{p})+\Pr[X_m=C_{m,0}]\,\tau_{C,0}(\boldsymbol{p})\right) \tag{28}$$

where $\tau_{m,0}$, $\tau_{C,a,0}$, *and* $\tau_{C,0}$ *are provided in Theorems 1–3, respectively. Here,* $\mathcal{A}_m$ *is the total probability of being associated with an access point when content* m *is requested, which is given by*

$$\mathcal{A}_m = \Pr[X_m = F_{m,0}] + \Pr[X_m = F_{a,0}] + \Pr[X_m = C_{m,0}] \tag{29}$$

Proof. Noting that the delay is conditioned on being associated with an access point. Thus, by total probability theorem, the probability is divided by the probability of the events space $\mathcal{A}_m$. □

To reduce the complexity of Theorem 8, the asymptotic average delay as $\frac{P}{N_0} \to \infty$ is considered in the following corollary.

Corollary 2 (Asymptotic delay)**.** *When* $\frac{P}{N_0} \to \infty$, *the average delay of* u_0 *can be expressed as*

$$\begin{aligned}
\tau_\infty(\boldsymbol{p}) &\triangleq \lim_{\frac{P}{N_0}\to\infty} \tau(\boldsymbol{p}) \\
&= T \sum_{m\in\mathcal{M}} \frac{a_m}{\mathcal{A}_m} \Bigg(\Big(\Pr[X_m = F_{m,0}] \times 2\pi p_m \lambda_F \int_0^R d \exp\Big(-\pi p_m \lambda_F \left(1 - \mathcal{U}(\boldsymbol{p})\right) d^2\Big) dd \Big) \\
&\quad + \Bigg(\Pr[X_m = F_{a,0}] \times \Big(2\pi \Lambda_m \lambda_F \int_0^R d \exp\Big(-\pi \Lambda_m \lambda_F (1 - \mathcal{V}(\boldsymbol{p}))\, d^2\Big) dd \\
&\qquad + 2\pi\lambda_C \int_0^\infty d \exp\Big(-\pi\lambda_C (1 - \mathcal{G})\, d^2\Big) dd \Big) \Bigg) \\
&\quad + \Big(\Pr[X_m = C_{m,0}] \times 2\pi\lambda_C \int_0^R d \exp\Big(-\pi\lambda_C (1 - \mathcal{G})\, d^2\Big) dd \Big) \Bigg)
\end{aligned} \tag{30}$$

Proof. The proof is analogous to the proof of Corollary 1. □

4. Problem Formulation and Optimization

It is noted that the STP and delay are fundamentally affected by the content caching distribution $\boldsymbol{p}$. Thus, the multi-objective problem to minimize the delay and maximize the STP (or equivalently minimize the negative STP) by optimizing the content caching distribution $\boldsymbol{p}$ can formulated as follows

Problem 1. *Multi-Objective Caching Optimization*

$$\begin{aligned}
&\min_{\boldsymbol{p}} \; -q(\boldsymbol{p}), \tau(\boldsymbol{p}) \\
&\textit{subjected to } (2), (3)
\end{aligned} \tag{31}$$

To reduce the computational complexity of Problem 1, we consider the asymptotic multi-objective optimization problem in the high SNR regime (i.e., $\frac{P}{N_0} \to \infty$).

Problem 2. *Asymptotic Multi-Objective Caching Optimization when* $\frac{P}{N_0} \to \infty$

$$\begin{aligned}
&\min_{\boldsymbol{p}} \; -q_\infty(\boldsymbol{p}), \tau_\infty(\boldsymbol{p}) \\
&\textit{subjected to } (2), (3)
\end{aligned} \tag{32}$$

Due to the complex forms of the STP and delay, their convexity cannot be assured. Thus, Problems 1 and 2 are generally non-convex.

An efficient algorithm to solve large-scale non-convex optimization problems is the cuckoo search algorithm (CSA) [21]. CSA was firstly proposed in [22] as a simple nature-inspired meta-heuristic global optimization algorithm. In CSA, the breeding behavior of cuckoo birds is mimicked by generating new nests (i.e., solutions) via Lévy flight and

random walk in each generation. In [23], CSA was modified to solve multi-objective optimization problems. CSA outperforms the other meta-heuristic global optimization algorithm in terms of simplicity, number of objective function evaluations, and execution time [24,25]. However, in constrained problems, the success rate of CSA that utilizes the penalty method is low, which is due to the feasibility of the optimal solution is not easy to be guaranteed when there are multiple constraints with different scales or equality constraints. Thus, to assure the feasibility of the optimal solution of MOCSA, we propose PMOCSA.

In the proposed PMOCSA outlined in Algorithm 1, the feasibility of the nests is assured by the projection optimization performed in steps 3, 7, and 14, which can be demonstrated as a search within the feasible set of nests for the closest nest to the projected nest $\breve{\boldsymbol{p}}^i$. Accordingly, the projection optimization problem is formulated as follows

Problem 3. *Projection Optimization*

$$\begin{aligned} &\min_{\boldsymbol{p}^i} \; ||\boldsymbol{p}^i - \breve{\boldsymbol{p}}^i||^2 \\ &\textit{subjected to } (2), (3) \end{aligned} \tag{33}$$

where $||.||$ *is the Euclidean norm. The optimal solution of the projection problem can be computed as* $p^i_m = \min\{\left[\breve{p}^i_m - \zeta\right]^+, 1\}$, $\forall m \in \mathcal{M}$, *where* ζ *satisfies* $\sum_{m\in\mathcal{M}} \min\{\left[\breve{p}^i_m - \zeta\right]^+, 1\} = 1$ *and* $[x]^+ \triangleq \max\{x, 0\}$.

The new nests generated by Lévy flight in step 6 can be obtained as follows

$$\tilde{\boldsymbol{p}}^i_{new} = \boldsymbol{p}^i + \epsilon \otimes \boldsymbol{L}(\Psi) \tag{34}$$

where $\tilde{\boldsymbol{p}}^i_{new}$ and $\boldsymbol{p}^i$ are the i-th new and current nest, respectively. ϵ denotes the step size scaling factor that is related to the scale of the problem, $\otimes$ is the entry-wise multiplication notation, $\Psi \in [0.3, 1.99]$ stands for the Lévy distribution index, and $\boldsymbol{L}(\Psi) = (L_m(\Psi))_{m\in\mathcal{M}}$ represents the Lévy vector, where the components of which can be calculated by Mantegna's algorithm as follows [26]

$$L_m(\Psi) = \frac{\Omega}{|\Theta|^{1/\Psi}}, \qquad \forall m \in \mathcal{M} \tag{35}$$

where $\Omega \overset{d}{\sim} \mathcal{N}(0, \sigma^2_\Omega)$ and $\Theta \overset{d}{\sim} \mathcal{N}(0, \sigma^2_\Theta)$ are random samples drawn from normal distribution of zero mean and variance σ^2_Ω and σ^2_Θ, respectively, where

$$\sigma^2_\Omega = \left[\frac{\sin(\pi\Psi/2)\,\Gamma(1+\Psi)}{\Psi\, 2^{(\Psi-1)/2}\,\Gamma((1+\Psi)/2)}\right]^{1/\Psi} \tag{36}$$

$$\sigma^2_\Theta = 1 \tag{37}$$

where $\Gamma(.)$ is the gamma function.

To encourage the localization of the search as the nest get closer to the solution, we consider a decreasing Lévy flight step size scaling factor ϵ, which is calculated at each iteration by

$$\epsilon(t) = \frac{N_c - t}{N_c} + \frac{1}{50\,M} \tag{38}$$

where t is the iteration index and N_c is the maximum number of iterations.

In step (13), a fraction of the nests is discarded by probability of P_a, then the same number of nests are regenerated by random walks as follows

$$\ddot{\boldsymbol{p}}^{*i} = \boldsymbol{p}^i + \delta\left(\boldsymbol{p}^k - \boldsymbol{p}^l\right) \tag{39}$$

where $\boldsymbol{p}^k$ and $\boldsymbol{p}^l$ are at two randomly selected nests, and δ is uniformly distributed random number in the interval (0,1).

Algorithm 1: Projected Multi-Objective Cuckoo Search Algorithm (PMOCSA).

1 set the maximum number of iterations N_C, the population size N_P, and the abandon probability P_a;
2 randomly generate initial population $\breve{\boldsymbol{p}}^i$, $(i = 1, 2, \cdots, N_P)$;
3 project $\breve{\boldsymbol{p}}$ onto the set of the variables satisfying (1) and (2) to obtain a feasible population $\boldsymbol{p}$;
4 evaluate the fitness values of each nest, i.e., $-q_\infty(\boldsymbol{p}^i)$ and $\tau_\infty(\boldsymbol{p}^i)$;
5 **while** $t \leq N_C$ **do**
6 randomly generate a new nest $\tilde{\boldsymbol{p}}^i_{new}$ via Lévy flight;
7 project $\tilde{\boldsymbol{p}}^i_{new}$ onto the set of the variables satisfying (1) and (2) to obtain a feasible nest $\boldsymbol{p}^i_{new}$;
8 evaluate the fitness values of $\boldsymbol{p}^i_{new}$ and check if it is Pareto optimal;
9 choose a nest $\boldsymbol{p}^j$ randomly among $\boldsymbol{p}$;
10 **if** *the fitness values of* $\boldsymbol{p}^i_{new}$ *dominate those of* $\boldsymbol{p}^j$ **then**
11 $\boldsymbol{p}^j \leftarrow \boldsymbol{p}^i_{new}$;
12 **end**
13 randomly abandon a fraction P_a of worse nests and build new ones, such that nest $\ddot{\boldsymbol{p}}^{*i}$ is generated via random walk;
14 project $\ddot{\boldsymbol{p}}^{*i}$ onto the set of the variables satisfying (1) and (2) to obtain a feasible nest $\ddot{\boldsymbol{p}}^i$;
15 **if** *the fitness values of* $\ddot{\boldsymbol{p}}^i$ *dominate those of* $\boldsymbol{p}^i$ **then**
16 $\boldsymbol{p}^i \leftarrow \ddot{\boldsymbol{p}}^i$;
17 **end**
18 sort the nests and find the current Pareto optimal solution;
19 **end**

In steps 10 and 15, the solution $\Big(-q_\infty(\boldsymbol{p}^i), \tau_\infty(\boldsymbol{p}^i) \Big)$ dominates solution $\Big(-q_\infty(\boldsymbol{p}^j), \tau_\infty(\boldsymbol{p}^j) \Big)$ if and only if $-q_\infty(\boldsymbol{p}^i) \leq -q_\infty(\boldsymbol{p}^j)$, $\tau_\infty(\boldsymbol{p}^i) \leq \tau_\infty(\boldsymbol{p}^j)$ and either or both $-q_\infty(\boldsymbol{p}^i) \neq -q_\infty(\boldsymbol{p}^j)$ and $\tau_\infty(\boldsymbol{p}^i) \neq \tau_\infty(\boldsymbol{p}^j)$. Thus, a solution is said to be non-dominated if no other solution dominates it.

It is noteworthy to highlight that there exist multiple solutions of a multi-objective optimization problem that are represented by Pareto front, which is defined as the set of non-dominated solutions.

Note that the optimal solution of the projection optimization problem can be obtained using *fmincon* function of MATLAB optimization tool box. Denote K as the maximum number of function evaluation in the projection optimization. Then, the time complexity in each iteration due to the projection optimization in steps 7 and 13 of PMOCSA is at most $\mathcal{O}(2KN_p)$. Whereas, the time complexity of the non-dominated sorting is $\mathcal{O}(n(2N_P)^2)$, where $n = 2$ is the number of objective functions. Thus, the time complexity of PMOCSA is $\mathcal{O}(N_C(2KN_p + 8N_P^2))$.

5. Numerical Results and Discussions

In this section, the numerical results of PMOCSA and the proposed asymptotic caching scheme are presented in Section 5.1. Then, the results are comprehensively summarized and discussed in Section 5.2.

5.1. Numerical Results

This subsection presents the performance evaluation of the proposed PMOCSA and the asymptotic multi-objective caching scheme. The performance of PMOCSA is compared

with the MOCSA proposed in [23] with a penalty factor of 1×10^6. Whereas, the performance of the proposed multi-objective caching scheme is compared with two well-known caching schemes. The first benchmark scheme is 'Popular' and proposed in [27], in which the F-APs cache only the most popular content. The Second is 'Uniform', which refers to the caching scheme proposed in [28], wherein the content are randomly cached with equal probabilities. It is assumed that the two benchmark schemes adopt the same wireless backhauling and association model of the proposed multi-objective caching scheme.

The parameters of the PMOCSA and MOSCA used in this paper are $N_C = 5000$, $N_P = 15$, $\Psi = 1.5$, and $P_a = 0.25$. In Figure 2, we plot the average Pareto fronts obtained after performing 50 trails and all the points visited by the optimization algorithm in the trails. It is observed that in general MOCSA fails to converge to a Pareto front. We can also observe that MOCSA visits a very small number of feasible points during the optimization. Whereas, the search within the feasible space due to the projection optimization in PMOCSA leads to the convergence to Pareto front. Figure 2 also demonstrates that the Pareto fronts obtained by PMOCSA achieves higher STPs and lower delays than the closest feasible points to the solution obtained by MOCSA.

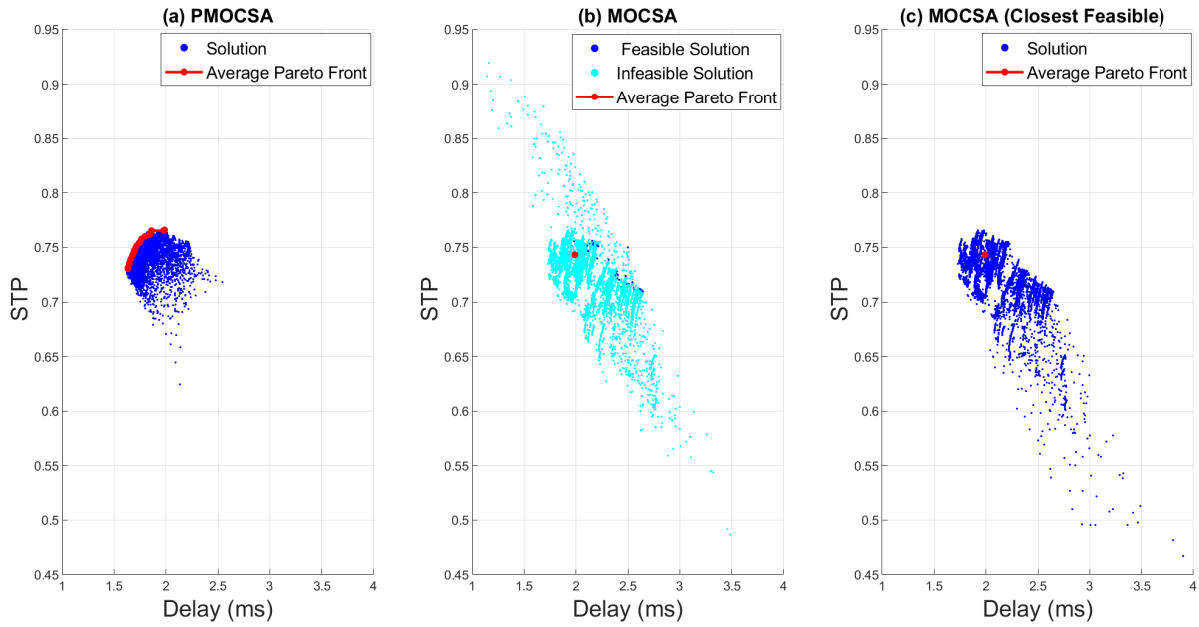

Figure 2. Euclidean distance versus iteration index at $M = 10$, $\lambda_U = 0.05$ users/m^2, $\lambda_F = 0.01$ fog access points (F-APs)/m^2, $\lambda_C = 0.001$ cloud access points (C-APs)/m^2, $R = 50$ m, $\gamma = 0.8$, $\alpha = 4$, $W_F = 10$ MHz, $W_C = 100$ MHz, $T = 1$ ms, and $\tau = 0.01$ Mbps.

To investigate the convergence rate of PMOCSA and MOCSA over all trails, we consider the average Euclidean distance between the Pareto fronts in consecutive iterations as a performance metric, i.e., the average sum of Euclidean distances between the corresponding non-dominated solutions on the Pareto fronts obtained in iteration t and $t-1$. Figure 3 shows that PMOCSA achieves lower average Euclidean distances and converges faster than MOCSA.

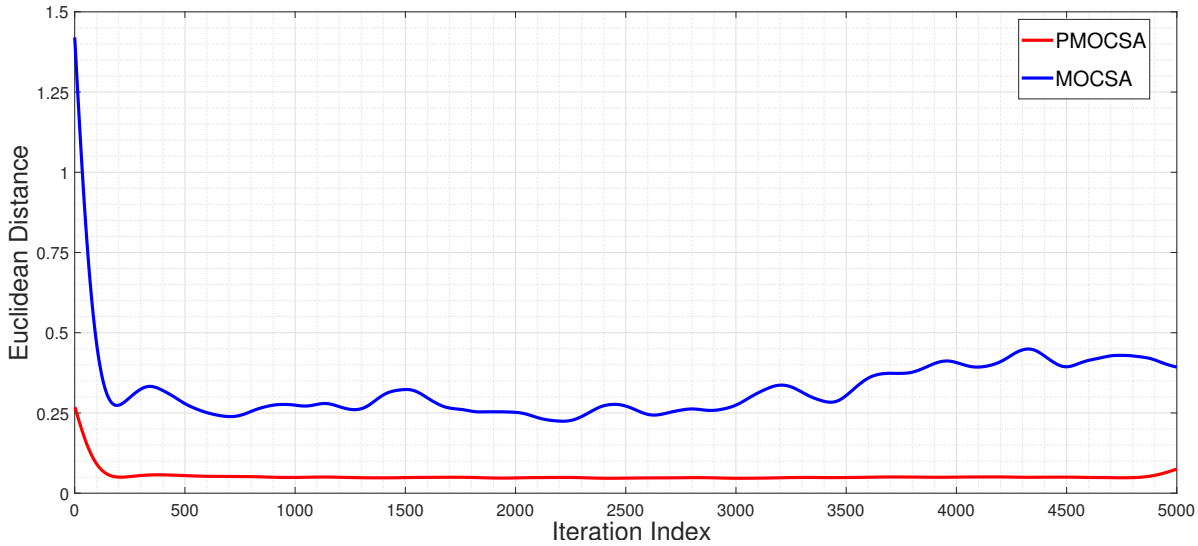

Figure 3. Successful transmission probability (STP) and delay values of all visited points during the optimization at $M = 10$, $\lambda_U = 0.05$ users/m^2, $\lambda_F = 0.01$ F-APs/m^2, $\lambda_C = 0.001$ C-APs/m^2, $R = 50$ m, $\gamma = 0.8$, $\alpha = 4$, $W_F = 10$ MHz, $W_C = 100$ MHz, $T = 1$ ms, and $\tau = 0.01$ Mbps.

The performance of the proposed asymptotic multi-objective caching scheme using PMOCSA is investigated for different network parameters in Figures 4–11. Specifically, Figure 4 plots the objective functions versus the discovery range R, where the surface represents the Pareto fronts and each black point represents a non-dominated optimal solution. Figure 4 demonstrates a growth in the STP with the discovery range for all schemes. This happens because of the higher probability with discovery range of u_0 being associated with an access point to serve its request. The figure shows that the increase in the STP is slight in the Popular scheme, which is due to the requests for the contents, except the most popular one, can only be served by transit F-APs or the C-APs. However, the Popular scheme achieves higher STPs than the Uniform scheme owing to the higher probability of transit F-AP, since the most popular content is only directly cached in the Popular scheme.

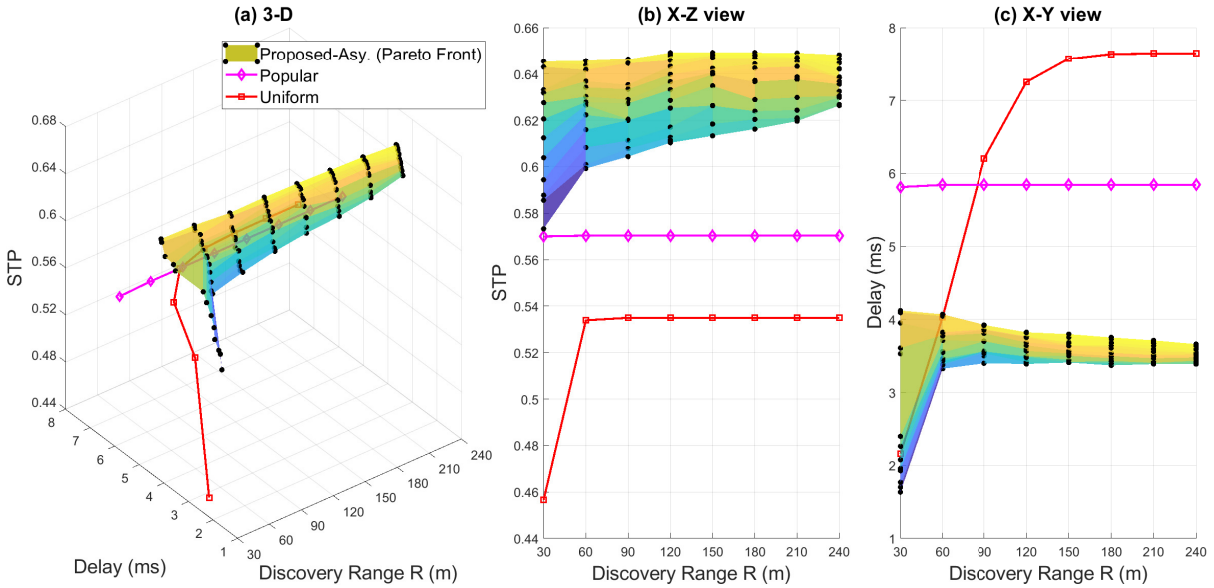

Figure 4. STP and delay versus discovery range at $M = 20$, $\lambda_U = 0.1$ users/m^2, $\lambda_F = 0.01$ F-APs/m^2, $\lambda_C = 0.001$ C-APs/m^2, $\gamma = 0.8$, $\alpha = 4$, $W_F = 10$ MHz, $W_C = 50$ MHz, $T = 1$ ms, and $\tau = 0.01$ Mbps.

Figure 4 also shows that the average delay increases with discovery range for all schemes, which is due to the high delay of the contents dissemination by the C-APs since the contents have higher probabilities with the discovery range of being fetched from the C-APs directly or via transit F-APs. The poor performance of the contents dissemination by the C-APs is also the reason of the high delays of the Popular scheme at low discovery ranges (i.e., $R < 90$ m), and the higher delays of Uniform scheme beyond this range due to the high probability of fetching the requested contents from the C-APs and the high separation distances between the u_0 and the transit F-APs. The Pareto fronts in Figure 4 indicates that the proposed multi-objective scheme can achieve higher STP, lower average delay, or both simultaneously than the benchmark schemes. We can observe that at high discovery ranges (i.e., $R > 90$ m) the proposed scheme achieves 5–15% higher STPs and 30–40% lower average delays than the Popular scheme. It is also observed that the low average delays are obtained at the low STPs. Thus, a trade-off between the STP and delay is required to balance the performance.

The relationship between the Zipf exponent γ, STP and delay is illustrated in Figure 5, wherein a growth in the STP with the Zipf exponent is shown for the Popular scheme, which is due to the cached most popular content has higher probability of being requested with increase in the Zipf exponent. This also results in lower average delay until the turning point of $\gamma = 1.2$. The high average delay beyond the turning point is due to the low probability of utilizing the F-APs as transit F-APs, which is owing to the very high probability of requesting the cached most popular content. Whereas, Figure 5 shows that the Zipf exponent has no impact on Uniform scheme since the contents are evenly cached at the F-APs. Figure 5 also demonstrates that the proposed scheme outperforms benchmark schemes, and the performance improvement increases with the Zipf exponent, where the obtained Pareto fronts achieve up to 40% higher STPs and 85% lower average delays than the benchmark schemes. It is also observed that the average delay of the proposed scheme has no turning point, i.e., the delay always decreases with the Zipf exponent, which is due to the optimized content placement.

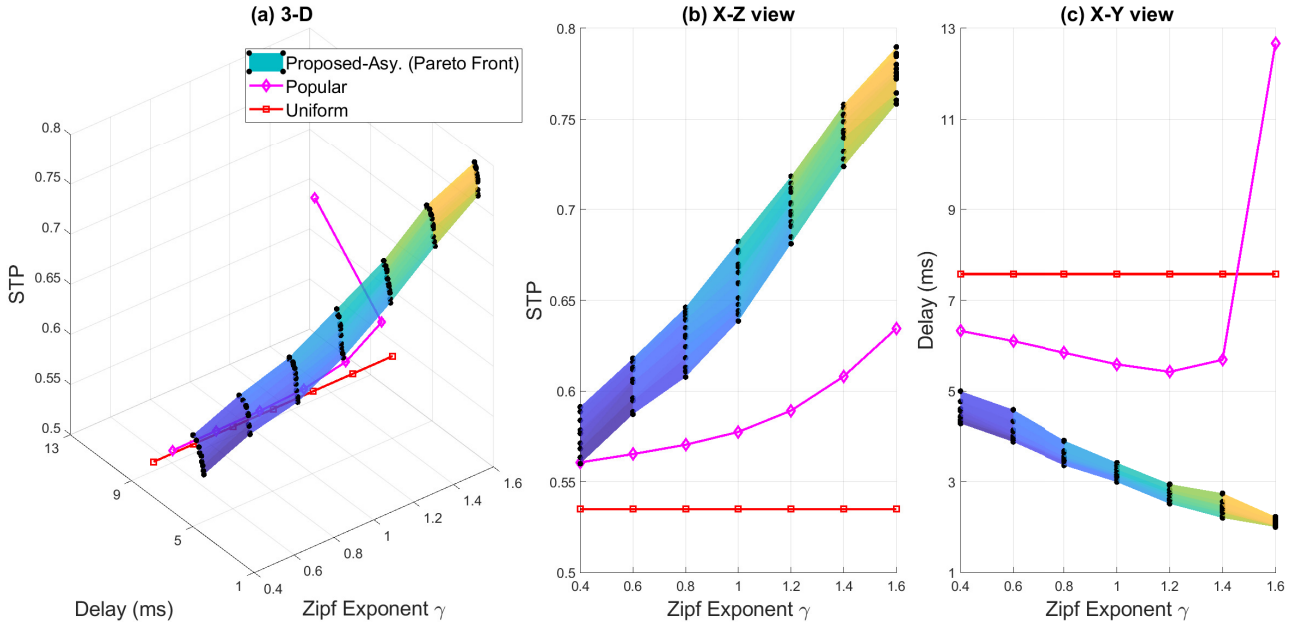

Figure 5. STP and delay versus Zipf exponent at $M = 20$, $\lambda_U = 0.1$ users/m^2, $\lambda_F = 0.01$ F-APs/m^2, $\lambda_C = 0.001$ C-APs/m^2, $R = 150$ m, $\alpha = 4$, $W_F = 10$ MHz, $W_C = 50$ MHz, $T = 1$ ms, and $\tau = 0.01$ Mbps.

Figure 6 plots the STP and delay versus the F-APs' bandwidth W_F. Figure 6 shows that the performance improves with the F-APs' bandwidth for all schemes. We can also observe that the impact of the F-APs' bandwidth on the Uniform scheme is higher than on the Popular scheme. This is due to the user's requests in the Uniform scheme are often

served by the direct F-APs. Figure 6 also shows that the proposed scheme performs better than the benchmark schemes.

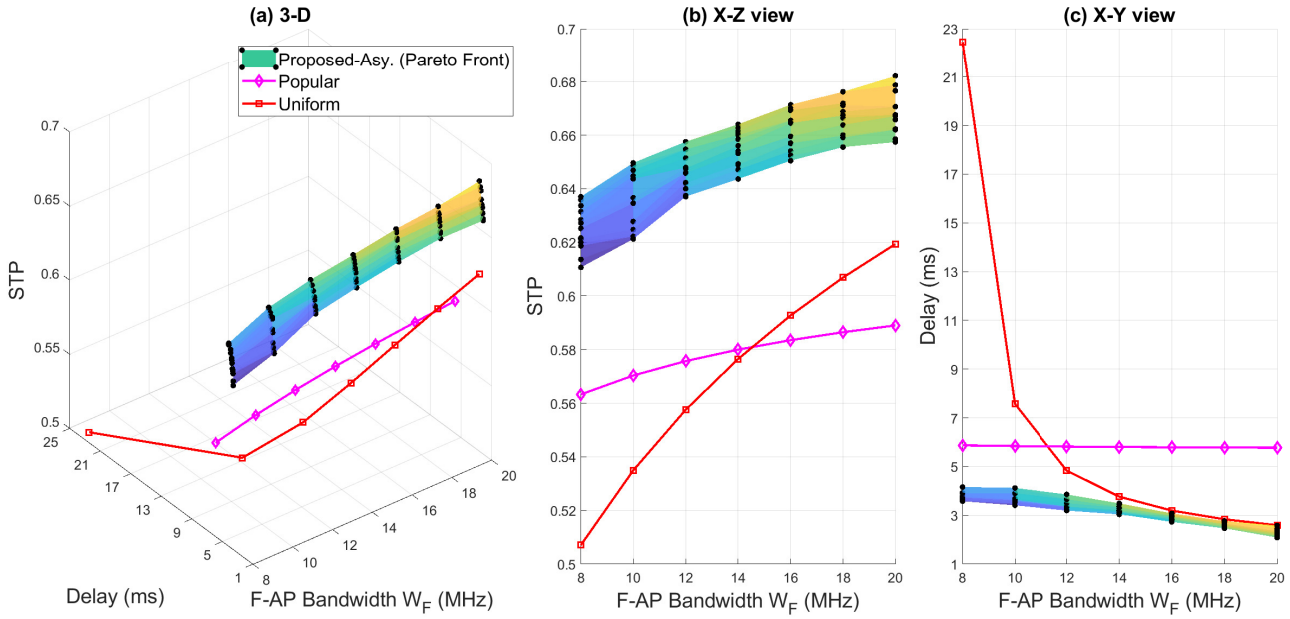

Figure 6. STP and delay versus F-APs' bandwidth at $M = 20$, $\lambda_U = 0.1$ users/m^2, $\lambda_F = 0.01$ F-APs/m^2, $\lambda_C = 0.001$ C-APs/m^2, $R = 150$ m, $\gamma = 0.8$, $\alpha = 4$, $W_C = 50$ MHz, $T = 1$ ms, and $\tau = 0.01$ Mbps.

In Figure 7, we plot the STP and delay versus the C-APs' bandwidth W_C. It can be observed that both STP and delay improves with increasing the C-APs' bandwidth for all schemes. However, the influence of the C-APs' bandwidth is higher on the Popular scheme as the requested contents expect the most popular are fetched from the C-APs directly or by the transit F-APs. It can also be observed the proposed scheme achieves higher STPs and lower delays than the benchmark schemes.

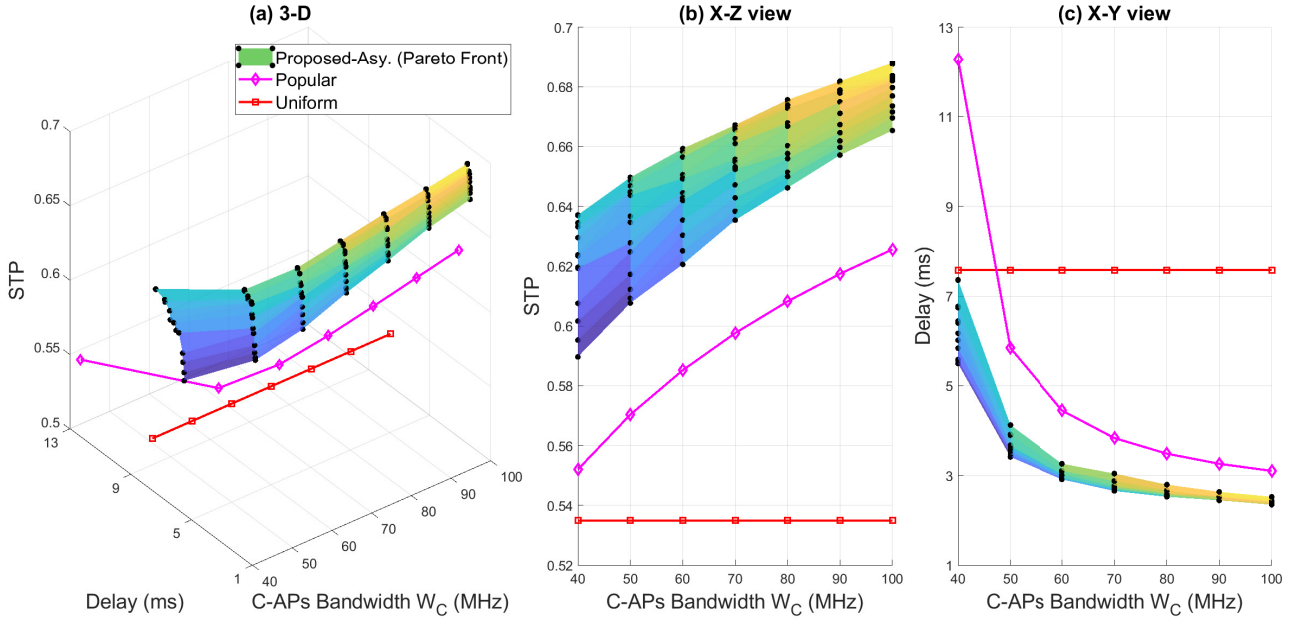

Figure 7. STP and delay versus C-APs' bandwidth at $M = 20$, $\lambda_U = 0.1$ users/m^2, $\lambda_F = 0.01$ F-APs/m^2, $\lambda_C = 0.001$ C-APs/m^2, $R = 150$ m, $\gamma = 0.8$, $\alpha = 4$, $W_F = 10$ MHz, $T = 1$ ms, and $\tau = 0.01$ Mbps.

Figure 8 illustrates the relationship between the total number of contents M, STP and delay. Figure 8 shows that the performance degrades for all schemes as the total number of contents increases, which can be explained by the increase in the total number of contents, results in caching smaller percentage of them within the discovery range and in a lower popularity of the most popular content. It is observed that the total number of contents has higher influence on the Uniform scheme than the other schemes, which is due to the fact that the users are often served by direct F-APs. Figure 8 also shows that the Uniform scheme performs worse than the Popular scheme with the increase in the total number of contents. This due to the lower probabilities of the direct and transit F-APs with total number of contents. The obtained Pareto fronts shown in Figure 8 demonstrates that the proposed scheme outperforms the benchmark schemes.

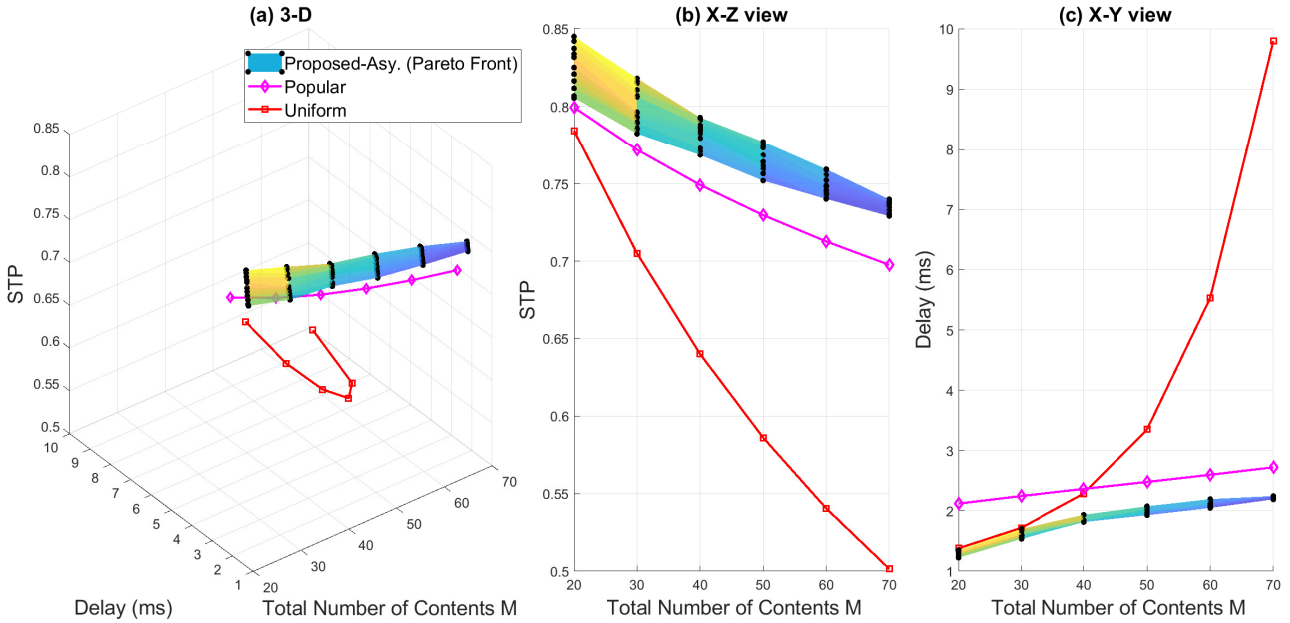

Figure 8. STP and delay versus total number of contents at $\lambda_U = 0.1$ users/m^2, $\lambda_F = 0.01$ F-APs/m^2, $\lambda_C = 0.001$ C-APs/m^2, $R = 150$ m, $\gamma = 0.8$, $\alpha = 4$, $W_F = 100$ MHz, $W_C = 500$ MHz, $T = 1$ ms, and $\tau = 0.01$ Mbps.

In Figure 9 we plot the STP and delay versus the user density λ_U. Figure 9 shows that the user density has no influence on the Uniform scheme owing to the requested contents are often downloaded from the direct F-APs. However, the increase in the number of users degrades the performance of the Popular scheme, wherein the contents are often served by transit F-APs, which is due to the decrease in the probability of using a F-AP for transit because there are more users requesting the most popular content. Whereas, the proposed scheme always performs better than the benchmark schemes as the cache placement is optimized to reduce the impact of the user density on the performance, and to provide a better utilization of the direct and transit F-APs.

Figure 10 shows the relationship between the F-AP density λ_F, STP, and the average delay. It is observed that the performance of Uniform scheme improves with F-AP density. This can be explained by, with the increase in the F-AP density, a higher number of the F-APs resides within the discovery range, which results in an increase of the probabilities of the direct and transit F-APs. However, in the Popular scheme, the STP and delay improve with the F-AP density until the turning points $\lambda_F = 8 \times 10^{-3}$ users/m^2 and $\lambda_F = 7 \times 10^{-3}$ users/m^2, respectively, then the performance degrades as the high interference originating from the F-APs dominates the gained improvement in the probability of transit F-APs. Figure 10 also shows that the proposed scheme outperforms the benchmark schemes, which is owing to the optimal utilization of the F-APs as direct or transit F-APs gained by optimizing the cache placement.

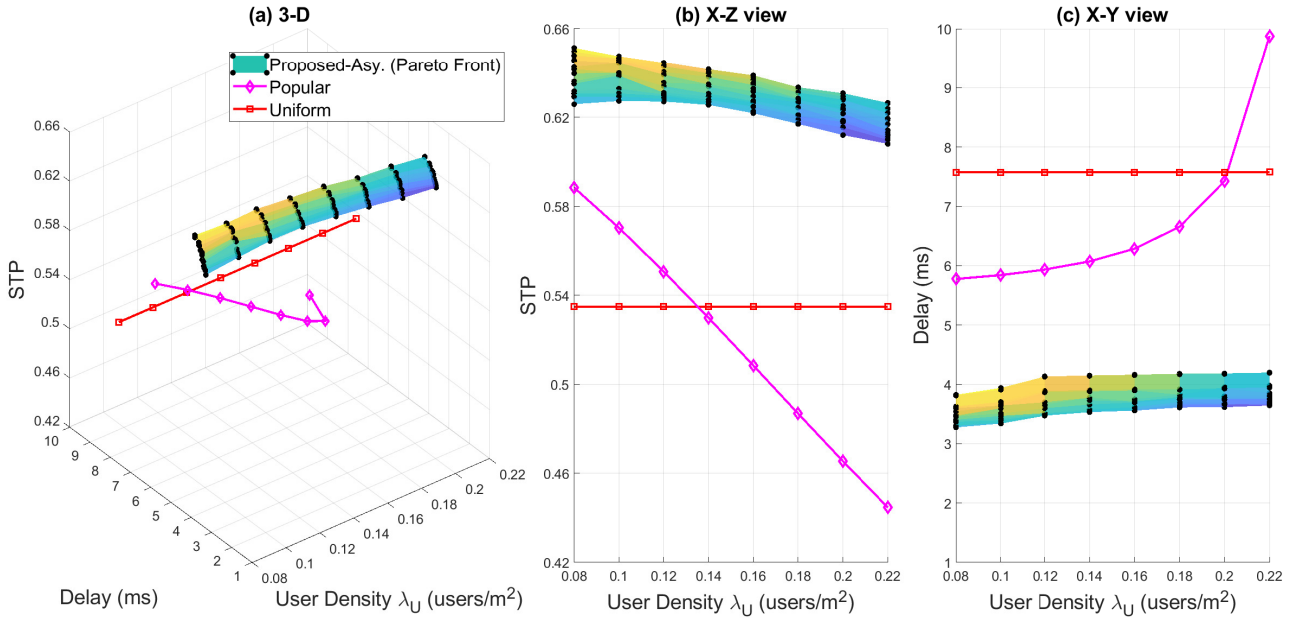

Figure 9. STP and delay versus user density at $M = 20$, $\lambda_F = 0.01$ F-APs/m^2, $\lambda_C = 0.001$ C-APs/m^2, $R = 150$ m, $\gamma = 0.8$, $\alpha = 4$, $W_F = 10$ MHz, $W_C = 50$ MHz, $T = 1$ ms, and $\tau = 0.01$ Mbps.

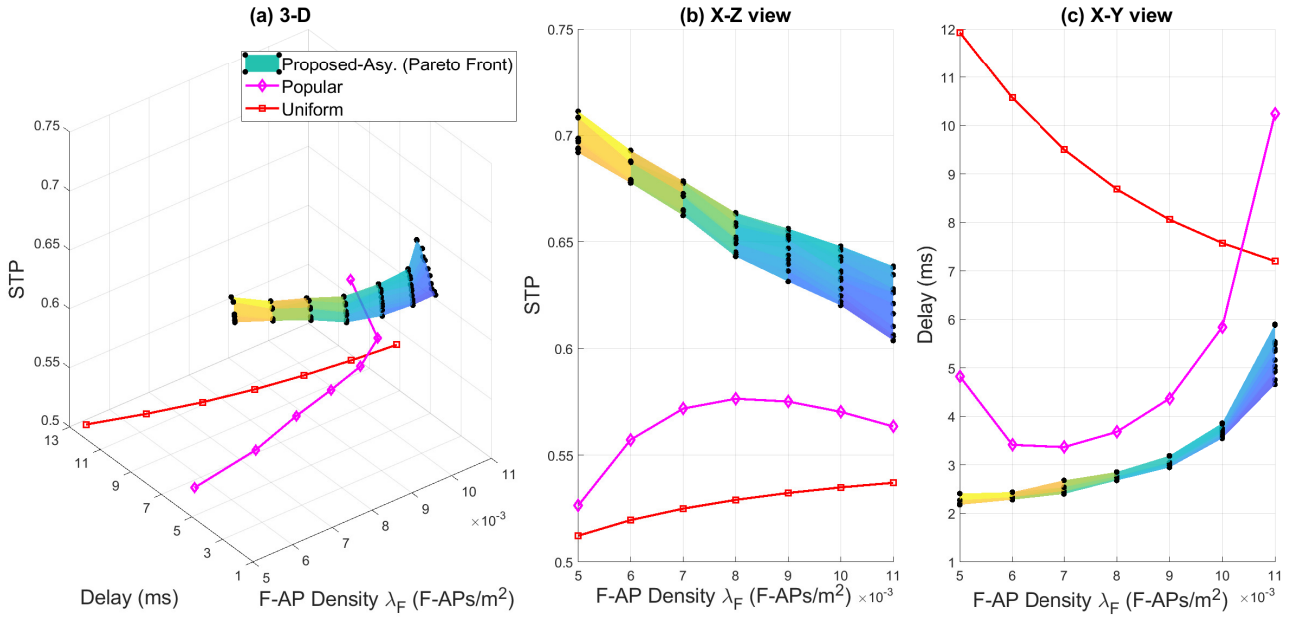

Figure 10. STP and delay versus F-AP density at $M = 20$, $\lambda_U = 0.1$ users/m^2, $\lambda_C = 0.001$ C-APs/m^2, $R = 150$ m, $\gamma = 0.8$, $\alpha = 4$, $W_F = 10$ MHz, $W_C = 50$ MHz, $T = 1$ ms, and $\tau = 0.01$ Mbps.

Figure 11 illustrates the impact of the C-AP density λ_C on the STP and average delay. Figure 11 demonstrates that the performance of the proposed scheme and the Popular scheme improves with the C-AP density as results of the improvement in the links between the C-APs to the transit F-APs and the higher probability of association directly with the C-APs to download the requested contents. Whereas, since the F-APs in the Uniform scheme are often working in direct mode, the increase in the C-AP density generates a high aggregated interference, which results in a performance degradation. Figure 11 also shows that the proposed multi-objective scheme always outperforms the Uniform and Poplar schemes.

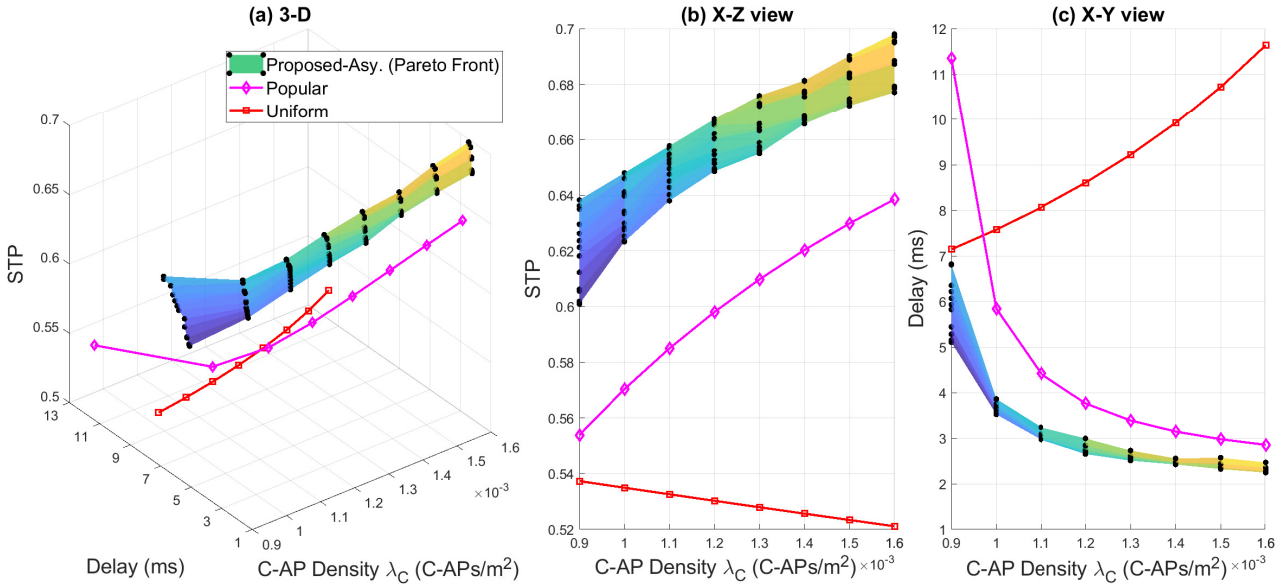

Figure 11. STP and delay versus C-AP density at $M = 20$, $\lambda_U = 0.1$ users/m^2, $\lambda_F = 0.01$ F-APs/m^2, $R = 150$ m, $\gamma = 0.8$, $\alpha = 4$, $W_F = 10$ MHz, $W_C = 50$ MHz, $T = 1$ ms, and $\tau = 0.01$ Mbps.

5.2. Discussions

In the previous subsection, it has been shown that the original MOCSA with well-known penalty method to handle the constraints generally fails to obtain a feasible Pareto front for Problem 2. Whereas, the proposed PMOCSA can efficiently converge to a feasible Pareto front, which represents the set of the non-dominated optimal solutions of the asymptotic multi-objective caching problem. The achieved superior performance of PMOCSA is not accompanied with extra hardware requirements, since PMOCSA can be implemented as simple code. Note that each solution on the obtained Pareto front using PMOCSA represents a distinct caching distribution. Therefore, an extra protocol might be required to distribute the contents to the F-APs according to the desired caching distribution.

Moreover, the performance of the proposed caching scheme using PMOCSA was evaluated for different network parameters and compared with two well-known caching schemes, i.e., the Popular and Uniform schemes, where the proposed scheme could always achieve higher performance, i.e., higher STPs and lower delays, than the benchmark schemes for all the studied network parameters. The numerical results showed that in the considered F-RAN system, the STP and delay are not always inversely related. In other words, the increase in the STP does not always result in a decrease in the delay, which is mainly due to the gained STP improvement by fetching the requested contents from the C-APs is accompanied with a large average delays as a result of the C-APs' poor content dissemination. The impacts of the network parameter on the performance of the proposed scheme were also analyzed. Specifically, the results demonstrated an increase in the STP and a decrease in the delay with the increase in the Zipf exponent, F-AP transmission bandwidth, C-AP transmission bandwidth, and C-AP density, and a decrease in the STP and an increase in the delay with the increase in the total number of contents, user density, and F-AP density. Whereas, the increase the discovery range led to an increase in the STP and delay, which is due to the higher probability of fetching the requested contents from the C-APs. Finally, bearing in mind that the Pareto front represents a set of solutions. Thus, a trade-off between the STP and delay is required to obtain the desired performance.

6. Conclusions

In this paper, we investigated the multi-objective proactive caching problem in wireless backhauled F-RAN. To formulate the optimization problem, first, we derived the expressions of the associations probability, STP, and delivery delay using stochastic geome-

try tools. To overcome the complexity of the STP and delay in the general SINR regime, we derived expressions of the asymptotic STP and delay in the high SNR regime. Then, we formulated the asymptotic multi-objective caching problem to minimize the delay or maximize the STP. Finally, we proposed a novel PMOCSA to obtain the Pareto front of the optimization problem. The numerical simulation results showed that PMOCSA outperforms MOCSA in terms of achieving a feasible Pareto front and the rate in which it converges. Finally, the proposed multi-objective caching scheme is shown to perform better than the well-known caching schemes.

Author Contributions: Conceptualization, A.B.-B., M.N.H., E.H., K.D., and T.F.T.M.N.I.; methodology, A.B.-B. and M.N.H.; software, A.B.-B. and M.N.H.; validation, K.D. and T.F.T.M.N.I.; formal analysis, A.B.-B. and M.N.H.; investigation, M.N.H. and K.D.; resources, M.N.H., E.H., and K.D.; data curation, A.B.-B. and M.N.H.; writing—original draft preparation, A.B.-B. and M.N.H.; writing—review and editing, E.H., K.D., and T.F.T.M.N.I.; visualization, A.B.-B., M.N.H., E.H., K.D., and T.F.T.M.N.I.; supervision, M.N.H., E.H., and K.D.; project administration, K.D. and T.F.T.M.N.I.; funding acquisition, K.D., E.H., and T.F.T.M.N.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research work has been supported by Ministry of Higher Education under the Fundamental Research Grant Scheme (FRGS/1/2020/TK0/UM/02/30).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funder had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A. Proof of Lemma 1

$\Pr[X_m = F_{m,0}]$ can be defined as the probability that there is at least one F-AP caching content m within R. Then, we have;

$$\begin{aligned}\Pr[X_m = F_{m,0}] &\triangleq \Pr[N(F_m) > 0] \\ &= 1 - \Pr[N(F_m) = 0] \\ &\overset{(a)}{=} 1 - \exp\left(-\pi p_m \lambda_F R^2\right)\end{aligned} \tag{A1}$$

where $N(F_m)$ is the number of F-APs caching content m within R. Here, (a) is obtained using the null property of PPP, i.e., $\Pr[N(F_m) = 0] = \exp\left(-\pi p_m \lambda_F R^2\right)$.

Appendix B. Proof of Lemma 2

An available F-AP with respect to content m is defined as a F-AP that is not caching content m and caches an inactive content, where the probability of caching the inactive content $\mu \in \mathcal{M} \setminus m$ is calculated as $p_\mu b_\mu$. Accordingly, the probability of the F-AP availability when content m is requested can be obtained by summing over the set of contents $\mathcal{M} \setminus m$ as follows

$$\begin{aligned}\Lambda_m &= Pr[\textit{ The F-AP caches inactive content } \neq m] \\ &= \sum_{\mu \in \mathcal{M} \setminus m} p_\mu b_\mu\end{aligned} \tag{A2}$$

Appendix C. Proof of Lemma 3

Noting that when u_0 requests content m, it is associated with a transit F-AP when there is no F-AP caching content m and there is at least one available F-AP within R. Then,

the probability of this event, i.e., $Pr[X_m = F_{a,0}]$, can be obtained using the proprieties of PPP as follows

$$\begin{aligned} Pr[X_m = F_{a,0}] &= Pr[N(F_m) = 0, N(F_a) > 0] \\ &\overset{(b)}{=} Pr[N(F_m) = 0]Pr[N(F_a) > 0] \\ &= Pr[N(F_m) = 0](1 - Pr[N(F_a) = 0]) \\ &\overset{(c)}{=} \exp\left(-\pi p_m \lambda_F R^2\right)\left(1 - \exp\left(-\pi \Lambda_m \lambda_F R^2\right)\right) \end{aligned} \tag{A3}$$

where $N(F_a)$ refers to the number of available F-APs. Equality (b) is obtained by noting that the events $N(F_m)$ and $N(F_a)$ are independent; (c) is obtained using the null property of PPPs.

Appendix D. Proof of Lemma 4

Since u_0 requesting content m is associated with the nearest C-AP C_0 if there is no F-AP caching m nor an available F-AP with respect to m are exist within R, and at least one C-AP exists within R, the probability of this event is given by

$$\begin{aligned} Pr[X_m = C_0] &= Pr[N(F_m) = 0, N(F_a) = 0, N(C) > 0] \\ &\overset{(d)}{=} Pr[N(F_m) = 0]Pr[N(F_a) = 0]Pr[N(C) > 0] \\ &= Pr[N(F_m) = 0]Pr[N(F_a) = 0](1 - Pr[N(C) = 0]) \\ &\overset{(e)}{=} \exp\left(-\pi p_m \lambda_F R^2\right)\exp\left(-\pi \Lambda_m \lambda_F R^2\right)\left(1 - \exp\left(-\pi \lambda_C R^2\right)\right) \\ &= \exp\left(-\pi (p_m + \Lambda_m) \lambda_F R^2\right)\left(1 - \exp\left(-\pi \lambda_C R^2\right)\right) \end{aligned} \tag{A4}$$

where $N(C)$ is the number of C-APs. (d) is due to the independence of the events $N(F_m)$, $N(F_a)$ and $N(C)$. The null property of PPP is used to obtain (e).

Appendix E. Proof of Theorem 1

Denote $I_m \triangleq \sum_{l \in \Phi_m \setminus F_{m,0}} D_{l,0}^{-\alpha} |h_{l,0}|^2$, $I_{-m} \triangleq \sum_{l \in \Phi_{-m}} D_{l,0}^{-\alpha} |h_{l,0}|^2$, and $I_C \triangleq \sum_{l \in \Phi_C} D_{l,0}^{-\alpha} |h_{l,0}|^2$ as the interference from the F-APs caching content m, the F-APs that are not caching content m, and the C-APs, respectively. Then, $q_{m,0}$ conditioned on the distance $D_{0,0} = d \in [0, R]$ can be calculated as follows

$$\begin{aligned} q_{m,0,D_{0,0}}(\boldsymbol{p}, d) &\triangleq \Pr[W_F \log_2(1 + SINR_{m,0}) \geq \xi | D_{0,0} = d] \\ &= E_{I_m, I_{-m}, I_C}\left[\Pr\left[|h_{0,0}|^2 \geq s\left(I_m + I_{-m} + I_C + \frac{N_0}{P}\right)\right]\right] \\ &\overset{(f)}{=} E_{I_m, I_{-m}, I_C}\left[\exp\left(-s\left(I_m + I_{-m} + I_C + \frac{N_0}{P}\right)\right)\right] \\ &\overset{(g)}{=} \underbrace{E_{I_m}[\exp(-sI_m)]}_{\triangleq \mathcal{L}_{I_m}(s,d)} \underbrace{E_{I_{-m}}[\exp(-sI_{-m})]}_{\triangleq \mathcal{L}_{I_{-m}}(s,d)} \underbrace{E_{I_C}[\exp(-sI_C)]}_{\triangleq \mathcal{L}_{I_C}(s,d)} \exp\left(-s\frac{N_0}{P}\right) \end{aligned} \tag{A5}$$

where $s = \left(2^{\frac{\xi}{W_F}} - 1\right) d^{\alpha}$, equality (f) is due to $|h|^2 \overset{d}{\sim} \exp(1)$, equality (g) is due to the independence of the PPPs and the independence of the Rayleigh fading channels, and $\mathcal{L}_{I_m}(s, d)$,

$\mathcal{L}_{I_{-m}}(s,d)$, and $\mathcal{L}_{I_C}(s,d)$ denote the Laplace transforms of I_m, I_{-m}, and I_C, respectively. $\mathcal{L}_{I_m}(s,d)$ can be calculated as follows

$$\begin{aligned}\mathcal{L}_{I_m}(s,d) &= E\left[\exp\left(-s\sum_{l\in\Phi_m\setminus F_{m,0}} D_{l,0}^{-\alpha}\,|h_{l,0}|^2\right)\right]\\ &= E\left[\prod_{l\in\Phi_m\setminus F_{m,0}} \exp\left(-sD_{l,0}^{-\alpha}\,|h_{l,0}|^2\right)\right]\\ &\overset{(h)}{=} \exp\left(-2\pi p_m\lambda_F\int_d^\infty\left(1-\frac{1}{1+s\,r^{-\alpha}}\right)r\,dr\right)\\ &\overset{(i)}{=} \exp\left(\frac{-2\pi}{\alpha}\,p_m\,\lambda_F\,s^{\frac{2}{\alpha}}\,\beta'\left(\frac{2}{\alpha},1-\frac{2}{\alpha},\frac{1}{1+s\,d^{-\alpha}}\right)\right)\\ &= \exp\left(\frac{-2\pi}{\alpha}\,p_m\,\lambda_F\left(2^{\frac{\xi}{W_F}}-1\right)^{\frac{2}{\alpha}}d^2\,\beta'\left(\frac{2}{\alpha},1-\frac{2}{\alpha},2^{\frac{-\xi}{W_F}}\right)\right)\end{aligned} \tag{A6}$$

here, (h) is obtained using the probability generating functional [29], and the following changes of variables $sr^{-1/\alpha}$ to t, after that $1/(1+t^{-\alpha})$ to w are used to obtain (i). Next, by noting that the density of Φ_{-m} is $(1-p_m)\lambda_F$ and following the same steps above, $\mathcal{L}_{I_{-m}}(s,d)$ can be calculated by

$$\mathcal{L}_{I_{-m}}(s,d) = \exp\left(\frac{-2\pi}{\alpha}\,(1-p_m)\,\lambda_F\left(2^{\frac{\xi}{W_F}}-1\right)^{\frac{2}{\alpha}}d^2\,\beta\left(\frac{2}{\alpha},1-\frac{2}{\alpha}\right)\right) \tag{A7}$$

In the same manner, $\mathcal{L}_{I_C}(s,d)$ can be given as

$$\mathcal{L}_{I_C}(s,d) = \exp\left(\frac{-2\pi}{\alpha}\,\lambda_C\left(2^{\frac{\xi}{W_F}}-1\right)^{\frac{2}{\alpha}}d^2\,\beta\left(\frac{2}{\alpha},1-\frac{2}{\alpha}\right)\right) \tag{A8}$$

Finally, $q_{m,0}(\boldsymbol{p})$ is calculated by removing the condition on the distance $D_{0,0}=d$ as follows

$$q_{m,0}(\boldsymbol{p}) = \int_0^R q_{m,0,D_{0,0}}(\boldsymbol{p},d)\,f_{D_{0,0}}(d)\mathrm{d}d \tag{A9}$$

here, $f_{D_{0,0}}(d) = 2\pi p_m\lambda_F d\exp\left(-\pi p_m\lambda_F d^2\right)$ which is obtained by noting that Φ_m is a homogeneous PPP with density $p_m\lambda_F$. Thus, we can prove Theorem (1).

Appendix F. Proof of Theorem 2

Let $I_a \triangleq \sum_{l\in\Phi_{a,m}\setminus F_{a,0}} D_{l,0}^{-\alpha}\,|h_{l,0}|^2$, $I_{-a} \triangleq \sum_{l\in\Phi_{-a,m}} D_{l,0}^{-\alpha}\,|h_{l,0}|^2$, $I_{C_0} \triangleq \sum_{l\in\Phi_C\setminus C_0} D_{l,a}^{-\alpha}\,|h_{l,a}|^2$, and $I_{F_a} \triangleq \sum_{l\in\Phi_F\setminus F_{a,0}} D_{l,a}^{-\alpha}\,|h_{l,a}|^2$ denote the interference from the available F-APs at u_0, the unavailable F-APs at u_0, the C-APs at $F_{a,0}$, and the F-APs at $F_{a,0}$, respectively. Next, the conditional STPs $q_{a,0,D_{a,0}}(\boldsymbol{p},d)$ and $q_{C,a,D_{C,a}}(\boldsymbol{p},d)$ conditioned on $D_{a,0}=d\in[0,R]$ and $D_{C,a}=d\in[0,\infty]$, respectively, can be obtained by following the same steps of (A5) as

$$q_{a,0,D_{a,0}}(\boldsymbol{p},d) = \mathcal{L}_{I_a}(s,d)\mathcal{L}_{I_{-a}}(s,d)\mathcal{L}_{I_C}(s,d)\exp\left(-s\frac{N_0}{P}\right) \tag{A10}$$

$$q_{C,a,D_{C,a}}(\boldsymbol{p},d) = \mathcal{L}_{I_{C_0}}(\tilde{s},d)\mathcal{L}_{I_{F_a}}(\tilde{s},d)\exp\left(-\tilde{s}\frac{N_0}{P}\right) \tag{A11}$$

here, $s = \left(2^{\frac{\xi}{W_F}} - 1\right) d^{\alpha}$ and $\tilde{s} = \left(2^{\frac{M\xi}{W_C}} - 1\right) d^{\alpha}$. Next, following the same procedure in (A6), we have

$$\mathcal{L}_{I_a}(s,d) = \exp\left(\frac{-2\pi}{\alpha}\Lambda_m \lambda_F \left(2^{\frac{\xi}{W_F}} - 1\right)^{\frac{2}{\alpha}} d^2 \beta'\left(\frac{2}{\alpha}, 1 - \frac{2}{\alpha}, 2^{\frac{-\xi}{W_F}}\right)\right) \tag{A12}$$

$$\mathcal{L}_{I_{-a}}(s,d) = \exp\left(\frac{-2\pi}{\alpha}(1-\Lambda_m) \lambda_F \left(2^{\frac{\xi}{W_F}} - 1\right)^{\frac{2}{\alpha}} d^2 \beta\left(\frac{2}{\alpha}, 1 - \frac{2}{\alpha}\right)\right) \tag{A13}$$

$$\mathcal{L}_{I_C}(s,d) = \exp\left(\frac{-2\pi}{\alpha}\lambda_C \left(2^{\frac{\xi}{W_F}} - 1\right)^{\frac{2}{\alpha}} d^2 \beta\left(\frac{2}{\alpha}, 1 - \frac{2}{\alpha}\right)\right) \tag{A14}$$

$$\mathcal{L}_{I_{C_0}}(\tilde{s},d) = \exp\left(\frac{-2\pi}{\alpha}\lambda_C \left(2^{\frac{M\xi}{W_C}} - 1\right)^{\frac{2}{\alpha}} d^2 \beta'\left(\frac{2}{\alpha}, 1 - \frac{2}{\alpha}, 2^{\frac{-M\xi}{W_C}}\right)\right) \tag{A15}$$

$$\mathcal{L}_{I_{F_a}}(\tilde{s},d) = \exp\left(\frac{-2\pi}{\alpha}\lambda_F \left(2^{\frac{M\xi}{W_C}} - 1\right)^{\frac{2}{\alpha}} d^2 \beta\left(\frac{2}{\alpha}, 1 - \frac{2}{\alpha}\right)\right) \tag{A16}$$

Finally, the PDF of $D_{a,0}$ and $D_{C,a}$ are given by $f_{D_{a,0}}(d) = 2\pi\Lambda_m\lambda_F\, d \times \exp\left(-\pi\Lambda_m\lambda_F d^2\right)$ and $f_{D_{C,a}}(d) = 2\pi\lambda_C\, d \exp\left(-\pi\lambda_C d^2\right)$, respectively. Thus, the condition on the distance can be removed as follows

$$q_{a,0}(\boldsymbol{p}) = \int_0^R q_{a,0,D_{a,0}}(\boldsymbol{p},d)\, f_{D_{a,0}}(d)\mathrm{d}d \tag{A17}$$

$$q_{C,a}(\boldsymbol{p}) = \int_0^{\infty} q_{C,a,D_{C,a}}(\boldsymbol{p},d)\, f_{D_{C,a}}(d)\mathrm{d}d \tag{A18}$$

Thus, we complete the proof.

Appendix G. Proof of Theorem 3

Let $I_{C_m} \triangleq \sum_{l\in\Phi_C\setminus C_{m,0}} D_{l,0}^{-\alpha}\,|h_{l,0}|^2$ and $I_F \triangleq \sum_{l\in\Phi_F} D_{l,0}^{-\alpha}\,|h_{l,0}|^2$ stands for the interference at u_0 originating from the C-APs except $C_{m,0}$ and the F-APs, respectively. Therefore, the conditional STP on the distance $D_{C,0} = d \in [0,R]$ can be expressed as

$$q_{C,0,D_{C,0}}(\boldsymbol{p},d) = \mathcal{L}_{I_{C_m}}(s,d)\mathcal{L}_{I_F}(s,d)\exp\left(-s\frac{N_0}{P}\right) \tag{A19}$$

where $s = \left(2^{\frac{M\xi}{W_C}} - 1\right) d^{\alpha}$. To proceed, the Laplace transforms of the interference can be obtained as

$$\mathcal{L}_{I_{C_m}}(s,d) = \exp\left(\frac{-2\pi}{\alpha}\lambda_C \left(2^{\frac{M\xi}{W_C}} - 1\right)^{\frac{2}{\alpha}} d^2 \beta'\left(\frac{2}{\alpha}, 1 - \frac{2}{\alpha}, 2^{\frac{-M\xi}{W_C}}\right)\right) \tag{A20}$$

$$\mathcal{L}_{I_F}(s,d) = \exp\left(\frac{-2\pi}{\alpha}\lambda_F \left(2^{\frac{M\xi}{W_C}} - 1\right)^{\frac{2}{\alpha}} d^2 \beta\left(\frac{2}{\alpha}, 1 - \frac{2}{\alpha}\right)\right) \tag{A21}$$

Then, the condition $D_{C,0} = d$ is removed to obtain $q_{C,0}$ as follows

$$q_{C,0}(\boldsymbol{p}) = \int_0^R q_{C,0,D_{C,0}}(\boldsymbol{p},d)\, f_{D_{C,0}}(d)\mathrm{d}d \tag{A22}$$

where the PDF of $D_{C,0}$ is given by $f_{D_{C,0}}(d) = 2\pi\lambda_C\, d \exp\left(-\pi\lambda_C d^2\right)$. Thus, we can prove Theorem 3.

References

1. Mukherjee, M.; Shu, L.; Wang, D. Survey of Fog Computing: Fundamental, Network Applications, and Research Challenges. *IEEE Commun. Surv. Tutor.* **2018**, *20*, 1826–1857. [CrossRef]
2. Habibi, M.A.; Nasimi, M.; Han, B.; Schotten, H.D. A Comprehensive Survey of RAN Architectures Toward 5G Mobile Communication System. *IEEE Access* **2019**, *7*, 70371–70421. [CrossRef]
3. Bani-Bakr, A.; Dimyati, K.; Hindia, M.N.; Wong, W.R.; Al-Omari, A.; Sambo, Y.A.; Imran, M.A. Optimizing the Number of Fog Nodes for Finite Fog Radio Access Networks under Multi-Slope Path Loss Model. *Electronics* **2020**, *9*, 2175. [CrossRef]
4. Emara, M.; Elsawy, H.; Sorour, S.; Al-Ghadhban, S.; Alouini, M.; Al-Naffouri, T.Y. Optimal Caching in 5G Networks With Opportunistic Spectrum Access. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 4447–4461. [CrossRef]
5. Peng, A.; Jiang, Y.; Bennis, M.; Zheng, F.; You, X. Performance Analysis and Caching Design in Fog Radio Access Networks. In Proceedings of the 2018 IEEE Globecom Workshops (GC Wkshps), Abu Dhabi, United Arab Emirates, 9–13 December 2018; pp. 1–6. [CrossRef]
6. Wang, R.; Li, R.; Wang, P.; Liu, E. Analysis and Optimization of Caching in Fog Radio Access Networks. *IEEE Trans. Veh. Technol.* **2019**, *68*, 8279–8283. [CrossRef]
7. Jiang, F.; Yuan, Z.; Sun, C.; Wang, J. Deep Q-Learning-Based Content Caching With Update Strategy for Fog Radio Access Networks. *IEEE Access* **2019**, *7*, 97505–97514. [CrossRef]
8. Jiang, Y.; Ma, M.; Bennis, M.; Zheng, F.; You, X. User Preference Learning-Based Edge Caching for Fog Radio Access Network. *IEEE Trans. Commun.* **2019**, *67*, 1268–1283. [CrossRef]
9. Jia, S.; Ai, Y.; Zhao, Z.; Peng, M.; Hu, C. Hierarchical content caching in fog radio access networks: Ergodic rate and transmit latency. *China Commun.* **2016**, *13*, 1–14. [CrossRef]
10. Liu, J.; Bai, B.; Zhang, J.; Letaief, K.B. Cache Placement in Fog-RANs: From Centralized to Distributed Algorithms. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 7039–7051. [CrossRef]
11. Wei, X. Joint Caching and Multicast for Wireless Fronthaulin Fog Radio Access Networks. In Proceedings of the 2017 IEEE 86th Vehicular Technology Conference (VTC-Fall), Toronto, ON, Canada, 24–27 September 2017; pp. 1–5. [CrossRef]
12. Li, Z.; Chen, J.; Zhang, Z. Socially Aware Caching in D2D Enabled Fog Radio Access Networks. *IEEE Access* **2019**, *7*, 84293–84303. [CrossRef]
13. Dang, T.; Peng, M. Joint Radio Communication, Caching, and Computing Design for Mobile Virtual Reality Delivery in Fog Radio Access Networks. *IEEE J. Sel. Areas Commun.* **2019**, *37*, 1594–1607. [CrossRef]
14. Wei, Y.; Yu, F.R.; Song, M.; Han, Z. Joint Optimization of Caching, Computing, and Radio Resources for Fog-Enabled IoT Using Natural Actor–Critic Deep Reinforcement Learning. *IEEE Internet Things J.* **2019**, *6*, 2061–2073. [CrossRef]
15. Jiang, Y.; Hu, Y.; Bennis, M.; Zheng, F.; You, X. A Mean Field Game-Based Distributed Edge Caching in Fog Radio Access Networks. *IEEE Trans. Commun.* **2020**, *68*, 1567–1580. [CrossRef]
16. Guo, B.; Zhang, X.; Sheng, Q.; Yang, H. Dueling Deep-Q-Network Based Delay-Aware Cache Update Policy for Mobile Users in Fog Radio Access Networks. *IEEE Access* **2020**, *8*, 7131–7141. [CrossRef]
17. Rahman, G.M.S.; Peng, M.; Yan, S.; Dang, T. Learning Based Joint Cache and Power Allocation in Fog Radio Access Networks. *IEEE Trans. Veh. Technol.* **2020**, *69*, 4401–4411. [CrossRef]
18. Jiang, Y.; Wan, C.; Tao, M.; Zheng, F.C.; Zhu, P.; Gao, X.; You, X. Analysis and Optimization of Fog Radio Access Networks With Hybrid Caching: Delay and Energy Efficiency. *IEEE Trans. Wirel. Commun.* **2021**, *20*, 69–82. [CrossRef]
19. Jiang, Y.; Peng, A.; Wan, C.; Cui, Y.; You, X.; Zheng, F.; Jin, S. Analysis and Optimization of Cache-Enabled Fog Radio Access Networks: Successful Transmission Probability, Fractional Offloaded Traffic and Delay. *IEEE Trans. Veh. Technol.* **2020**, *69*, 5219–5231. [CrossRef]
20. Yu, S.M.; Kim, S. Downlink capacity and base station density in cellular networks. In Proceedings of the 2013 11th International Symposium and Workshops on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), Tsukuba, Japan, 13–17 May 2013; pp. 119–124.
21. Vo, D.N.; Schegner, P.; Ongsakul, W. Cuckoo search algorithm for non-convex economic dispatch. *IET Gener. Transm. Distrib.* **2013**, *7*, 645–654. [CrossRef]
22. Yang, X.; Deb, S. Cuckoo Search via Lévy flights. In Proceedings of the 2009 World Congress on Nature Biologically Inspired Computing (NaBIC), Coimbatore, India, 9–11 December 2009; pp. 210–214. [CrossRef]
23. Yang, X.S.; Deb, S. Multiobjective cuckoo search for design optimization. *Comput. Oper. Res.* **2013**, *40*, 1616–1624. [CrossRef]
24. Khodier, M. Comprehensive study of linear antenna array optimisation using the cuckoo search algorithm. *IET Microw. Antennas Propag.* **2019**, *13*, 1325–1333. [CrossRef]
25. Nguyen, T.T.; Vo, D.N. The application of one rank cuckoo search algorithm for solving economic load dispatch problems. *Appl. Soft Comput.* **2015**, *37*, 763–773. [CrossRef]
26. Mantegna, R.N. Fast, accurate algorithm for numerical simulation of Lévy stable stochastic processes. *Phys. Rev. E* **1994**, *49*, 4677–4683. [CrossRef]
27. Baştuğ, E.; Bennis, M.; Kountouris, M.; Debbah, M. Cache-enabled small cell networks: Modeling and tradeoffs. *EURASIP J. Wirel. Commun. Netw.* **2015**, *2015*, 1–11. [CrossRef] [PubMed]

28. Tamoor-ul-Hassan, S.; Bennis, M.; Nardelli, P.H.J.; Latva-Aho, M. Modeling and analysis of content caching in wireless small cell networks. In Proceedings of the 2015 International Symposium on Wireless Communication Systems (ISWCS), Brussels, Belgium, 25–28 August 2015; pp. 765–769. [CrossRef]
29. Haenggi, M.; Ganti, R. Interference in Large Wireless Networks. *Found. Trends Netw.* **2009**, *3*, 127–248. [CrossRef]

Short Biography of Authors

Alaa Bani-Bakr received B.Sc. and M.Sc. degrees in electrical/telecommunication engineering from Mutah University, Karak, Jordan, in 2002 and 2006, respectively. He is currently pursuing a Ph.D. with the University of Malaya, Kuala Lumpur, Malaysia. He was a Lecturer at the Department of Electrical Engineering, AlBaha University, Saudi Arabia, from 2010 to 2012. His research interests are fog radio access networks, cache-enabled wireless networks, mmWave communication systems, stochastic analysis, and optimization.

MHD Nour Hindia received a Ph.D. from the Faculty of Engineering in Telecommunication, University of Malaya, Kuala Lumpur, Malaysia, in 2015. He is currently involved with research in the field of wireless communications, especially in channel sounding, network planning, converge estimation, handover, scheduling, and quality of service enhancement for 5G networks. He is currently a Post-Doctoral Fellow from the Faculty of Engineering in Telecommunication, University of Malaya. Besides that, he is involved with research with the Research Group in Modulation and Coding Scheme for Internet of Things for Future Network. He has authored or co-authored a number of science citation index journals and conference papers. Dr. Hindia has participated as a Reviewer and a committee member of a number of ISI journals and conferences.

Kaharudin Dimyati graduated from the University of Malaya, Malaysia, in 1992. He received a Ph.D. from the University of Wales Swansea, U.K., in 1996. He is currently a Professor at the Department of Electrical Engineering, Faculty of Engineering, University of Malaya. Since joining the university, he is actively involved in teaching, postgraduate supervision, research, and also administration. To date, he has supervised 15 Ph.D. students and 32 masters by research students. He has published over 100 journal articles. He is a member of IET and IEICE. He is a Professional Engineer and a Chartered Engineer.

Effariza Hanafi received a B.Eng. in telecommunications (first class Hons.) from the University of Adelaide, Adelaide, S.A., Australia, and the Ph.D. degree in electrical and electronic engineering from the University of Canterbury, Christchurch, New Zealand, in 2010 and 2014, respectively. She joined the University of Malaya, Kuala Lumpur, Malaysia, where she is now a Senior Lecturer at the Faculty of Engineering. In 2015, she was the recipient of the University of Malaya Excellence Awards. She is currently a Senior Member for IEEE (Institute of Electrical and Electronics Engineers). Her main research interests include wireless communications, Internet of Things, cognitive radio, cooperative communications, 5G networks, and beyond.

Tengku Faiz Tengku Mohmed Noor Izam received a Ph.D. in electronic engineering from the University of Surrey, U.K., in 2016. He is currently a Lecturer with the Department of Electrical Engineering, University of Malaya, Malaysia. His research interests include, parasitic antenna, and MIMO system with antenna selection.