CrossMark

Fog radio access network system control scheme based on the embedded game model

Sungwook Kim

Abstract

As a promising paradigm for the 5G wireless communication system, a new evolution of the cloud radio access networks has been proposed, named as fog radio access networks (F-RANs). It is an advanced socially aware mobile networking architecture to provide a high spectral and energy efficiency while reducing backhaul burden. In particular, F-RANs take full advantages of social information and edge computing to efficiently alleviate the end-to-end latency. Based on the benefit of edge and cloud processing, key issues of F-RAN technique are radio resource allocation, caching, and service admission control. In this paper, we develop a novel F-RAN system control scheme based on the embedded game model. In the proposed scheme, spectrum allocation, cache placement, and service admission algorithms are jointly designed to maximize system efficiency. By developing a new embedded game methodology, our approach can capture the dynamics of F-RAN system and effectively compromises the centralized optimality with decentralized distribution intelligence for the faster and less complex decision making process. Through simulations, we compare the performance of our scheme to the existing studies and show how we can achieve a better performance under dynamic F-RAN system environments.

Keywords: Fog radio access network, Service admission control, Cache placement, Embedded game model, Radio resource allocation

1 Introduction

In the past decade, the evolution toward 5G is featured by the explosive growth of traffic in the wireless network, due to the exponentially increased number of user devices. Compared to the 4G communication system, the 5G system should bring billions of user devices into wireless networks to demand high bandwidth connections. Therefore, system capacity and energy efficiency should be improved to get the great success of 5G communications. Cloud radio access network (C-RAN) is an emerging architecture for the 5G wireless system. A key advantage of C-RAN is the possibility to perform cooperative transmissions across multiple edge nodes for the centralized cloud processing. However, the cloud processing comes at the cost of the potentially large delay entailed by fronthaul transmissions. It may become a major performance bottleneck of a C-RAN system per critical indicators such as spectral efficiency and latency [1–3].

As an extension of C-RAN paradigm, fog computing is a promising solution to the mission critical tasks involving quick decision making and fast response. It is a distributed paradigm that provides cloud-like services to the network edge nodes. Instead of using the remoted cloud center, the fog-computing technique leverages computing resources at the edge of networks based on the decentralized transmission strategies. Therefore, it can help overcome the resource contention and increasing latency. Due to the effective coordination of geographically distributed edge nodes, the fog-computing approach can meet the 5G application constraints, i.e., location awareness, low latency, and supports for mobility or geographical distribution of services. The most frequently referred use cases for the fog-computing concept are related to the Internet of Things (IoT) [4, 5].

Correspondence: swkim01@sogang.ac.kr
Department of Computer Science, Sogang University, 35 Baekbeom-ro
(Sinsu-dong), Mapo-gu 121-742, Seoul, South Korea

Taking full advantages of fog computing and C-RANs, fog radio access networks (F-RANs) have been proposed as an advanced socially aware mobile networking architecture in 5G systems. F-RANs harness the benefits of, and the synergies between, fog computing and C-RAN in order to accommodate the broad range of Quality of Service (QoS) requirements of 5G mobile broadband communications [2]. In the F-RAN architecture, edge nodes may be endowed with caching capabilities to serve the local data requests of popular content with low latency. At the same time, a central cloud processor allocates radio and computational resources to each individual edge nodes while ensuring as much as various applications [6]. To maximize the F-RAN system performance, application request scheduling, cache placement, and communication resource allocation should be jointly designed. However, it is an extremely challenging issue.

In the architecture of F-RANs, multiple interest relevant system agents exist; they are the central cloud server (CS), edge nodes (ENs), and mobile users (MUs). The CS provides contents to download and allocates radio and communication resources to ENs. ENs, known as fog-computing-based access points (F-APs), manage the allocated radio resource and admit MUs to provide application services. MUs wish to enjoy different QoS services from the F-RAN system. Different system agents have their own benefits, but their benefits could conflict with each other, and each agent only cares about its own profit. Therefore, it is necessary to analyze the interactions among these conflicting system agents and design proper solutions. Although dozens of techniques have been proposed, a systematic study on the interactions among CS, F-APs, and MUs is still lacking [7].

Recently, game theory has been widely concerned and studied in solving conflict and cooperation in distributed optimization problems. As a branch of mathematics, it is suitable for analyzing the performance of multi-agent systems. In game theory, a variety of interactive behaviors can be described by different payoff functions and the outcome of a game depends on the combinations of the decisions taken by each agent. Depending on the way the decisions are taken, it is possible to classify the games as cooperative or non-cooperative games. Nowadays, game theory has been widely recognized as an important tool in many fields. In the last few years, it has attracted considerable attentions and has been investigated extensively in computer science and tele-communications [8]. This is one of the reasons why game theory is applied more and more intensively to cloud computing and network management fields.

However, the traditional game theoretic analysis should rely on the perfect information and idealistic behavior assumptions. Therefore, there is a quite general consensus to say that the predicted game solutions are useful but would be rarely observed in real world situations. Recently, specialized sub-branches of game theory have been developed to encounter this problem. The main goal of this study is also to develop a new game paradigm. To design a practical game model for the F-RAN system management, we adopt an online dynamic approach based on the interactive relationship among system agents. Our approach exploits a partial information on the game and obtains an effective solution under mild and practical assumptions. From the standpoint of algorithm designers, our approach can be dynamically implemented in the real-world F-RAN environments.

Motivated by the above discussion, we propose a novel F-RAN control scheme based on the methodology of game theory. In this study, we design a new game model, called embedded game, to effectively solve the conflict problem among F-RAN system agents. From the realistic point of view, we need not the complete knowledge of the system information. Instead, our game procedure imitates the interactive sequential game process while ensuring the system practicality. To summarize, the major contribution of this study is to provide a new game-based F-RAN control algorithm. Main features of our proposed game model are as follows: (i) the adjustable dynamics considering the current F-RAN system environment, (ii) an interactive online process based on the embedded game model, (iii) a cooperative control manner with the nested non-cooperative approach, (iv) the joint design to obtain synergistic and complementary features, and (v) a practicality under realistic system operation scenarios. Especially, the important novelty of our proposed scheme is obtained from the key principle of embedded game approach, which can better capture the reality of F-RAN operations. To the best of our knowledge, very little research has been done, and there is still few published work to discuss the joint F-RAN control algorithms.

This article is organized as follows. In the next section, we review some related schemes and their problems. In Section 3, we define the embedded game model considered in this paper and explain in detail the proposed F-RAN control scheme. In particular, this section provides fresh insights into the benefits and design of developed control algorithms. For convenience, the main steps of the proposed scheme are also listed in Section 3. Section 4 reports the simulation results and performance analysis. Lastly, we give our conclusion in Section 5. Open issues and challenges are also discussed in Section V.

2 Related work

Over the years, a lot of state-of-the-art research work on the radio access network control problem has been conducted. In [9], K. Sundaresan et al. proposed a scalable,

light-weight scheme for realizing the full potential of C-RAN systems. For small cells, this scheme determined configurations that maximized the traffic demand while simultaneously optimizing the compute resource usage in the baseband processing unit pool. Briefly, the developed scheme in [9] adopted a two-step approach: (i) the first-step determined the optimal combination of configurations; it needed to support the traffic demand from a set of small cells, and (ii) the second step consolidated the configurations to further reduce the compute resource usage [9].

The article [10] provided a brief overview of the infrastructure and logic structure of C-RAN systems. In addition, a new coordinated user scheduling algorithm and parallel optimum precoding algorithm were specifically designed based on the concept of a service cloud and a three-layer logical structure. This approach utilized extensive computation resources to improve the C-RAN system performance. Compared to traditional C-RAN algorithms, the developed scheme in [10] matched well with the C-RAN architecture while managing interference efficiently and accelerating the cooperation processing in parallel.

Q. Vien et al. [11–13] proposed a non-orthogonal multiple access-based power allocation scheme for C-RAN systems. In this scheme, base stations were allocated with different power levels depending on their distances to the cloud, and the optimal number of BSs in C-RAN systems was found to achieve an improved performance. Specifically, a successive interference cancellation mechanism was designed at the cloud to lay multiple base stations over each other in the power domain. Taking into account the constraints of the total available power and the cloud-edge throughput requirement, this approach has shown to support a higher number of base stations when compared to the existing scheme [11–13].

The paper [14] surveyed heterogeneous C-RAN research achievements and challenges, and provided a summary of recent advancements in the computing convergence of heterogeneous wireless networks. In particular, the system architecture, performance analysis, cloud-computing-based cooperative processing, networking techniques, key large-scale cooperative processing and networking techniques, including cloud-computing-based coordinated multipoint transmission, cooperative radio resource management, and self-organizing networks have been briefly summarized. Furthermore, potential challenges and open issues in heterogeneous C-RANs have been discussed as well [14].

In [15], Sengupta et al. provided a latency-centric analysis of the degrees of freedom of an F-RAN by accounting for the total content delivery delay across the fronthaul and wireless segments of the network. The main goal of the analysis was the identification of optimal caching, fronthaul, and edge transmission policies. In this study, authors detailed a general model and a novel performance metric, referred to as Normalized Delivery Time (NDT), which captured the worst-case delivery latency with respect to an ideal interference-free system. Finally, they revealed optimal caching-fronthaul transmission policies as a function of the system resources [15].

Azimi, et al. [16] considered an online caching setup, in which the set of popular files was time-varying and both cache replenishment and content delivery could take place in each time slot. They developed online caching and delivery schemes based on both reactive and proactive caching principles, and bounds on the corresponding achievable long-term NDTs were derived. In particular, a lower bound on the achievable long-term NDT was obtained. Using this bound, the performance loss caused by the variations in the set of popular files in terms of delivery latency was quantified by comparing the NDTs achievable under offline and online caching. Finally, numerical results were provided in which the performance of reactive and proactive online caching schemes were compared with offline caching [16].

The *Traffic Balancing and Dynamic Clustering* (*TBDC*) scheme investigated the joint design of multicast beamforming, dynamic clustering, and backhaul traffic balancing [17]. To minimize the power consumption for higher energy efficiency, the *TBDC* scheme designed the beamforming vectors and clustering pattern in the downlink of F-RAN. This approach balanced the backhaul traffic according to individual backhaul capacities, guaranteed the QoS of each user, and minimized the power consumption. Especially, the *TBDC* scheme dynamically excluded a radio unit from a cluster when it contributed comparatively less to the corresponding multicast group. If a radio unit contributed comparatively more to the corresponding multicast group, it would be involved in a cluster in order to guarantee the required QoS [17].

The *Cloud Structure with Edge Caching* (*CSEC*) scheme presented an information-theoretic model of F-RANs [6]. This scheme aimed at providing a latency-centric understanding of the degrees of freedom in the F-RAN network by accounting for the available limited resources in terms of fronthaul capacity, cache storage sizes, as well as power and bandwidth on the wireless channel. In addition, a new performance measure was introduced; it captured the worst-case latency incurred over the fronthaul. Finally, the *CSEC* scheme characterized the trade-off between the fronthaul and caching resources of the system while revealing optimal caching-fronthaul transmission policies [6].

The *Joint Optimization of Cloud and Edge* (*JOCE*) scheme introduced the joint design of cloud and edge processing for the downlink of F-RAN [3]. To design the delivery phase for an arbitrary pre-fetching strategy, transfer modes can be categorized into two classes: hard-transfer mode and soft-transfer mode. In the hard-transfer mode, non-cached files are communicated over the fronthaul links to a subset of access points. Therefore, this approach transfers hard information of subfiles that were not cached. In the soft-transfer mode, the fronthaul links are used to convey quantized baseband signals as in a cloud RAN. Therefore, this approach transfers a quantized version of the precoded signals for the missing files in line with the C-RAN paradigm. In the *JOCE* scheme, a novel superposition coding approach was proposed that was based on the hybrid use of the fronthaul links in both hard-transfer and soft-transfer modes. The problem of maximizing the delivery rate was tackled under fronthaul capacity and per enhanced remote radio head power constraints. This study was concluded that the *JOCE* scheme based on the superposition coding provided a more effective way and could have the potential to strictly outperform both conventional soft-transfer and hard-transfer modes [3].

2.1 Comparison and main contributions

Some earlier studies [9–16] have attracted considerable attention while introducing unique challenges in handling the cloud radio control problems. Even though these existing schemes dynamically control the cloud radio access network for the efficient system management, there are some difficulties to compare performance between these work with our proposed scheme. The scheme in [9] was developed only for small cells, i.e., houses, based on the partially centralized C-RAN model. The studies in [11–13] strongly concentrated on the non-orthogonal multiple access method to improve the spectral efficiency. Therefore, these studies were specially focused on the wireless downlink control problems in C-RAN systems. The papers [10, 14] surveyed various C-RAN research achievements and challenges and discussed issues of system architectures, spectral and energy efficiency performance, and promising key techniques. However, these surveys only covered research fields of traditional cloud radio access methods. In particular, the earlier studies [9–14] did not concern the issue of fog-computing paradigm. Therefore, they did not provide cloud-like services to the network edge nodes.

The studies [15, 16] considered the edge processing in F-RANs and specially investigated fundamental information theoretic limits. However, these schemes relied upon the cache-aided fog network paradigm while causing the extra cost to implement control mechanisms. This architecture-oriented approach was inappropriate

to fairly compare performance under general F-RAN system operations. The schemes in [3, 6, 17] have attracted considerable attention while introducing unique challenges in handling the edge cloud control problems. In this paper, we demonstrate through extensive simulation and analysis that our proposed scheme significantly outperforms these existing *TBDC*, *CSEC*, *JOCE* schemes.

The specific difference between the proposed scheme and the existing schemes in [3, 6, 9–17] is the decision making procedure; we designed a new embedded game model for the F-RAN system. Based on the step-by-step interactive mechanism, the proposed scheme jointly develops spectrum allocation algorithm and service admission algorithm; they are interlocked and serially correlated to capture the dynamics of F-RAN systems. Therefore, our approach is suitable in dynamically changing F-RAN environments and provides a well-balanced system performance than existing schemes, which were designed as one-sided protocols.

3 Embedded game model for F-RAN control algorithms

In recent years, the F-RAN system has attracted much attention due to its significant benefits to meet the enormous 5G application demands. Based on the general F-RAN architecture, different solutions have been proposed. In this section, the architecture of F-RAN is firstly introduced, and then the embedded game model is defined for effective F-RAN operations. Finally, we explain in detail about the proposed algorithm in the ten-step procedures.

A. Embedded game model for F-RAN systems

In the C-RAN architecture, all control functions and application storage are centralized at the CS, which requires a lot of MUs to transmit and exchange their data fast enough through the fronthaul link. To overcome this C-RAN's disadvantage with the fronthaul constraints, much attention has been paid to mobile fog computing and the edge cloud. The design of a fog-computing platform has been introduced to deliver large-scale latency-sensitive applications. To implement the fog-computing architecture, traditional edge nodes are evolved to the fog-computing-based access point (F-AP) by being equipped with a certain caching, cooperative radio resource, and computation power capability [2, 18]. The main difference between the C-RAN and the F-CRAN is that centralized storage cloud and control cloud functions are distributed to individual F-APs. Usually, F-APs are used to forward and process the received data, and interface to the CS through the fronthaul links. To avoid all traffic being loaded

directly to the centralized CS, some local traffic should be delivered from the caching located in F-APs. Therefore, each F-AP integrates not only the front radio spectrum but also the locally distributed cached contents and computation capacity. This approach can save the spectral usage of constrained fronthauls while decreasing the transmission delay. In conclusion, the main characteristics of F-RAN include ubiquity, decentralized management, and cooperation [2, 18]. The general architecture of a F-RAN system is shown in Fig. 1.

During the F-RAN system operations, system agents, i.e., CS, F-APs, MUs, should make decisions individually by considering the mutual-interaction relationship. Under the dynamic F-RAN environments, system agents try to maximize their own profits in a competitive or cooperative manner. In this study, we develop a new game model, called embedded game, for the F-RAN system. According to the decision making method, the embedded game procedure can be divided two phases. At the first phase, the CS and F-APs play a superordinated game; the CS distribute the available spectrum resource to each F-AP by using a cooperative manner. At the second phase, F-APs and MUs play subordinated games. By employing a non-cooperative manner, an individual F-AP selectively admits its corresponding MUs to provide different application services. Taken as a whole, multiple subordinated games are nested in the superordinated game.

Formally, we define the embedded game model $\mathbb{G} = \left\{ \mathbb{G}^{super}, \mathbb{G}^{sub}_{i,1 \leq i \leq n} \right\}$ where $\mathbb{G}^{super}$ is a superordinated game to formulate interactions between CS and F-APs, and $\mathbb{G}^{sub}_i$ is a subordinated
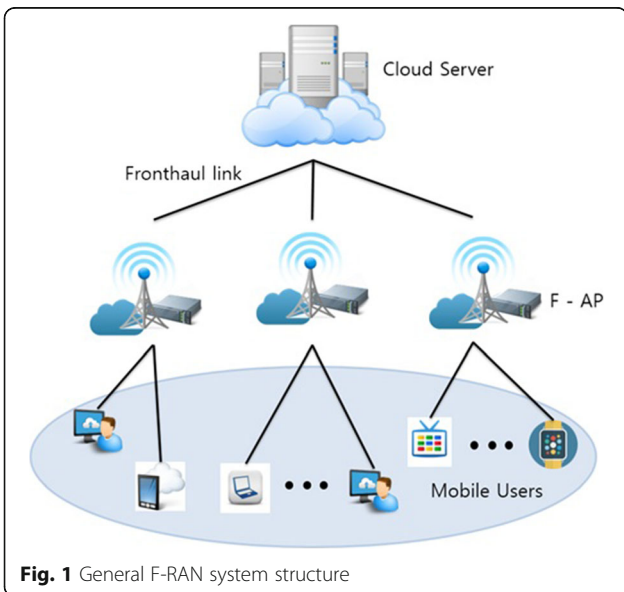
game to formulate interactions between the ith F-AP and its corresponding MUs. Firstly, the $\mathbb{G}^{super}$ can be defined as $\mathbb{G}^{super} = \left\{ \mathbb{N}, \mathcal{R}_{CS}, \boldsymbol{S}^{\mathcal{R}}_{CS}, U_{i,1 \leq i \leq n}, T \right\}$ at each time period t of gameplay.

- $\mathbb{N}$ is the finite set of $\mathbb{G}^{super}$ game players $\mathbb{N} = \{CS, F-AP_1, F-AP_2...F-AP_n\}$ where the total $n + 1$ number of $\mathbb{G}^{super}$ players; one CS and n F-APs.
- The total spectrum resources of CS is $\mathcal{R}_{CS}$, which would be distributed to n F-APs.
- $\boldsymbol{S}^{\mathcal{R}}_{CS} = \{\delta_1, \delta_2,.... \delta_n\}$ is the sets of CS's strategies for the spectrum resource allocation. δ_t in $\boldsymbol{S}^{\mathcal{R}}_{CS}$ is the allocated spectrum amount for the F - $AP_{t,1 \leq i \leq n}$.
- The $U_{t,\,1 \leq i \leq n}$ is the payoff received by the F - AP_t. It is estimated as the obtained outcome minus the cost from the spectrum resource allocation.
- The T is a time period. The $\mathbb{G}^{super}$ is repeated $t \in T < \infty$ time periods with imperfect information.

Secondly, the $\mathbb{G}^{sub}_i$ is the ith subordinated game, and it can be defined as $\mathbb{G}^{sub}_i = \left\{ \mathbb{M}_i, \boldsymbol{\Re}_i, \boldsymbol{S}^{\delta i}_{F-AP_i}, \boldsymbol{S}^{C^i}_{F-AP_i}, \boldsymbol{S}^{\sigma^i}_{F-AP_i}, \mathcal{U}^i_{j,1 \leq j \leq m}, T \right\}$ at each time period t of gameplay.

- $\mathbb{M}_i$ is the finite set of $\mathbb{G}^{sub}_i$ game players $\mathbb{M}_i = \{$F - $AP_i, MU^i_1,..., MU^i_m\}$ where $MU^i_{j,1 \leq j \leq m}$ is the jth MU in the area covered by the F - AP_i.
- The set of F - AP_i's resources is $\boldsymbol{\Re}_i = \{\delta_i, \mathcal{C}_i, \sigma_i\}$ where $\delta_i, \mathcal{C}_i, \sigma_i$ are the allocated spectrum resource, the computation capacity, and the placed cache files in the F - AP_i, respectively.
- $\boldsymbol{S}^{\delta_i}_{F-AP_i}, \boldsymbol{S}^{\mathcal{C}_i}_{F-AP_i}$ and $\boldsymbol{S}^{\sigma_i}_{F-AP_i}$ are the sets of F - AP_i's strategies for the spectrum allocation for MUs, the computation capacity assignment for MUs, and cache placement in the F - AP_i, respectively.
- The $\mathcal{U}^i_{j,1 \leq j \leq m}$ is the MU^{i}_j's payoff received by the F - AP_i.
- The T is a time period. The $\mathbb{G}^{sub}_i$ is repeated $t \in T < \infty$ time periods with imperfect information.

Table 1 lists the notations used in this paper.

B. Solution concept for the superordinated game

In the superordinated game, game players are CS and F-APs, and they are rational to reach a win-win situation. In many situations, each rational agent is able to improve its objectives without preventing others from improving their objectives. Therefore, they are more prone to coordinate and willing to play cooperative games [19]. Usually, solution concepts are different in different games. For the CS and F-AP interactions, the *Kalai and Smorodinsky Bargaining Solution* (KSBS) is an interesting solution concept. Like as the well-known *Nash Bargaining*

Fig. 1 General F-RAN system structure

Table 1 Parameters used in the proposed algorithm

Notations	Explanation
CS	Cloud server
F-APs	Fog-computing-based access points
MUs	Mobile users
ENs	Edge nodes
$\mathbb{N}$	The finite set of superordinated game players
$\mathscr{R}_{CS}$	The total spectrum resources of CS
δ_i	The allocated spectrum amount for the F‑AP$_i$
$\delta_i(t-\Delta t)$	The δ_i value at the time period $[t-\Delta t]$.
$U_i(\Delta t)$	The payoff received by the F‑AP$_i$ during the recent Δt time period
$\mathbb{M}_i$	The finite set of subordinated game players
$\mathfrak{R}_i$	The set of F‑AP$_i$'s resources
C_i	The computation capacity in the F‑AP$_i$
σ_i	The placed cache files in the F‑AP$_i$
U_j^i	The MU$_j^i$'s payoff received by the F‑AP$_i$
β	The parameter weighs the past experience by considering a trust decay over time
φ	The parameter specifies the impact of past experience
$T_i(t)$	At time t, the F‑AP$_i$'s trust assessment
F_{KSBS}^t	KSBS at time t
$d=(d_1,..d_n)$	Disagreement point when players cannot reach an agreement
ω_i^t	The player F‑AP$_i$'s bargaining power at time t
$\mathbb{R}^n$	A jointly feasible utility solution set
τ	Factor to characterize the file popularity
$\mathbb{M}=\{\mathscr{M}_1..\mathscr{M}_L\}$	A multimedia file set consists of L popular multimedia files
$\mathcal{Q}=[\mathscr{M}_1,...,\mathscr{M}_L]$	Vector to represent the popularity distribution among $\mathbb{M}$
$\mathbb{I}=[0,1]^{n\times L}$	A two-dimensional matrix to indicate the caching placement
$\mathcal{Z}_i^l$	The revenue from the lth file caching in the F‑AP$_i$,
$\mathfrak{C}_i^l$	The cost from the lth file caching in the F‑AP$_i$,
Θ_j^i	New service request of MU$_j^i$
$Min\_S(\Theta_j^i)$	The minimum spectrum requirement of Θ_j^i
$Min\_C(\Theta_j^i)$	The minimum computation requirement of Θ_j^i
χ^j	The currently using spectrum amount in the F‑AP$_i$
y^j	The currently using computation amount in the F‑AP$_i$
$\mathfrak{x}^i$	The current fronthaul transmission rate
$\mathfrak{M}^i$	The maximum fronthaul transmission rate

Solution (NBS), the KSBS also provides a fair and optimal solution in a cooperative manner. In addition, the KSBS can be used when the

feasible payoff set is not convex. It is the main advantage of KSBS over the NBS. Due to this appealing property, the KSBS approach has been practically implemented to solve real-world problems [8].

In order to show the effectiveness of the KSBS, it is necessary to evaluate each player's credibility. In this paper, we obtain the KSBS based on the F-APs' trustworthiness. This information can be inferred implicitly from the F-APs' outcome records. Therefore, we can enhance the effectiveness of KSBS while restricting the socially uncooperative F-APs. At time t, the F‑AP$_i$'s trust assessment $(\mathcal{T}_i(t))$ for the spectrum allocation process is denoted by

$$\mathcal{T}_i(t) = \{(1-\beta) \times \mathcal{T}_i(t-\Delta t)\}$$
$$+ \left\{ \beta \times \left(\left(\frac{U_i(\Delta t)}{\sum_{j=1}^n U_j(\Delta t)} \right) \Bigg/ \left(\frac{\delta_i(t-\Delta t)}{\mathscr{R}_{CS}} \right) \right) \right\} \tag{1}$$

$$\text{s.t., } \beta = \frac{(\phi \times \mathcal{T}_i(t-\Delta t))}{(1+\{\phi \times \mathcal{T}_i(t-\Delta t)\})} \text{ and } \phi \geq 0$$

where $U_i(\Delta t)$ is the throughput of the F‑AP$_i$ during the recent Δt time period, and $\delta_i(t-\Delta t)$ is the δ_i value at the time period $[t-\Delta t]$. The parameter β is used to weigh the past experience by considering a trust decay over time. In addition, we introduce another parameter   to specify the impact of past experience on $\mathcal{T}_i(t-\Delta t)$. Essentially, the contribution of current information increases proportionally as   increases. In this case, we can effectively adapt to the currently changing conditions while improving resiliency against credibility fluctuations [20]. In Eq. 1, the first term means the $\mathcal{T}_i$ value of the previous time period, and the second term represents the change ratio of δ_i to U_i at the current time period. In the point view of CS, $\mathcal{T}_i(t)$ is a weighted average of these two terms.

Under the dynamic F-RAN environment, we assume that F-APs request individually their spectrum resources to the CS at each time period. To adaptively respond to the current F-RAN system conditions, the sequential KSBS bargaining approach gets the different KSBS at each time period. It can adapt the timely dynamic F-RAN situations. At time t, the timed

$$\mathfrak{F}_{KSBS}^t\left(S_{cs}^{\mathscr{R}}\right) = \{\delta_1(t), \delta_2(t), ...\delta_n(t)\} = \left(\frac{\sup\{U_1^t(\delta_1(t))\}-d_1}{\omega_1^t \times \left(\mathbb{O}_1^t - d_1\right)} = ... = \frac{\sup\{U_i^t(\delta_i(t))\}-d_i}{\omega_i^t \times \left(\mathbb{O}_i^t - d_i\right)} = ... = \frac{\sup\{U_n^t(\delta_n(t))\}-d_n}{\omega_n^t \times \left(\mathbb{O}_n^t - d_n\right)}\right)$$

$$s.t., , = \mathbb{O}_i^t = \max\left\{U_i^t(\delta_i(t)) \mid U_i^t(\delta_i(t)) \in \mathbb{R}^n\right\}, \quad \omega_i^t = \frac{T_{i(t)}/}{\sum_{j=1}^n T_{j(t)}}$$

$$\text{and} \quad \sup\{U_i^t(\delta_i(t))\} = \sup\{U_i^t(\delta_i(t)) : \{(U_1^t(\delta_1(t)), ..., U_n^t(\delta_n(t)))\} \subset \mathbb{R}^n\} \tag{2}$$

KSBS ($\mathfrak{F}_{KSBS}^t$) for the spectrum resource problem is mathematically defined as; where $U_i^t(\delta_i(t))$ is the F - AP$_i$'s payoff with the strategy δ_i during the recent time period (Δt). $\mathbb{R}^n$ is a jointly feasible utility solution set, and a disagreement point (d) is an action vector $d = (d_1, .. d_n) \in \mathbb{R}^n$ that is expected to be the result if players, i.e., F-APs, cannot reach an agreement (i.e., zero in the system). ω_i^t ($0 < \omega_i^t < 1$) is the player F - AP$_i$'s bargaining power at time t, which is the relative ability to exert influence over other players. $\mathbb{O}_i^t$ is the ideal point of player F - AP$_i$ at time t. Therefore, players choose the best outcome subject to the condition that their proportional part of the excess over the disagreement is relative to the proportion of the excess of their ideal gains. Geometrically, the $\mathfrak{F}_{KSBS}^t$ $\left(S_{cs}^{\mathscr{R}}\right)$ is the intersection point between the bargaining set $S_{cs}^{\mathscr{R}}$ and the line, which is drawn from the disagreement point (d) to the best utilities, i.e., the ideal gains, of players. Simply, we can think that the KSBS is the maximal point which maintains the ratios of gains [21]. Therefore, $\mathfrak{F}_{KSBS}^t\left(S_{cs}^{\mathscr{R}}\right) = \{\delta_1(t), \delta_2(t), ... \delta_n(t)\} = \{\sup\{U_1^t(\delta_1(t))\}, \sup\{U_2^t(\delta_2(t))\}, ..., \sup\{U_n^t(\delta_n(t))\}\}$ is a joint strategy, which is taken by the CS at time t.

In non-deterministic settings, $\mathfrak{F}_{KSBS}^t\left(S_{cs}^{\mathscr{R}}\right)$ is a selection function to define a specific spectrum allocation strategy for every F-APs. Due to the main feature of KSBS, the increasing of bargaining set size in a direction favorable to a specific F-AP always benefits that F-AP. Therefore, in our superordinated game, self-interested F-AP can be satisfied during the F-RAN system operations. To practically obtain the $\mathfrak{F}_{KSBS}^t\left(S_{cs}^{\mathscr{R}}\right)$ in Eq. 2, we can re-think the KSBS as a weighted max-min solution like as

$$\mathfrak{F}_{KSBS}^t\left(S_{cs}^{\mathscr{R}}\right) = \{\delta_1(t), \delta_2(t), ...\delta_n(t)\} =$$
$$= arg \max_{\{\delta_1(t), \delta_2(t), ...\delta_n(t)\}} \left\{\min_{\delta_{i,1 \le i \le n}(t)} \left(\frac{\sup\{U_i^t(\delta_i(t))\}-d_i}{\omega_i^t \times \left(\mathbb{O}_i^t - d_i\right)}\right)\right\} \tag{3}$$

C. Solution concept for the subordinated games

Edge processing is a key emerging trend in the F-RAN system. It refers to the localization of computing, communication, and storage resources at the F-APs. In the F-RAN architecture, F-APs are connected to the CS through fronthaul links. Under this centralized structure, the performance of F-RANs is clearly constrained by the fronthaul link capacity; it incurs a high burden on fronthaul links. Therefore, a prerequisite requirement for the centralized CS processing is the high bandwidth and low latency fronthaul interconnections. However, during the operation of F-RAN system, unexpected growth of service requests may create a traffic congestion. It has a significant impact on the F-RAN performance. To overcome the disadvantages of F-RAN architecture imposed by the fronthaul constraints, new techniques have been introduced with the aim of reducing the delivery latency by limiting the need to communicate between the CS and MUs [6]. Currently, there are evidences that MUs' downloading of on-demand multimedia data is the major reason for the data avalanche over F-RAN; numerous repetitive requests on the same data lead to redundant transmissions. Usually, multimedia data are located in the CS and far away from MUs. To ensure an excellent QoS provisioning, an efficient solution is to locally store these frequently accessed data into the cache memory of F-APs while reducing the transmission latency; it is known as caching. This approach can effectively mitigate the unnecessary fronthaul overhead caused by MUs' repetitive service requests. Therefore, CS, F-APs and MUs are all the beneficiaries from the local caching mechanism [22]. In the subordinated game, an efficient caching mechanism is designed by carefully considering the relations and interactions among CS, F-APs and MUs. This approach can relive the heavy traffic load at fronthaul links, and also decrease the request latency; it results in better QoS [6]. A practical caching mechanism is coupled with the data placement. In our F-RAN architecture, we assume that a multimedia file set $\mathbb{M} = \{\mathcal{M}_1, ..., \mathcal{M}_L\}$ consists of L popular multimedia files in the CS, and files in $\mathbb{M}$ can be possibly cached in each F-AP. The popularity

distribution among $\mathbb{M}$ is represented by a vector $\mathcal{Q}$ = $[g_1,..., g_L]$. Generally, the vector $\mathcal{Q}$ can be modeled by a Zipf distribution [22];

$$g_l = \left. \left(\frac{1}{l^\tau}\right) \middle/ \left(\sum\nolimits_{f=1}^{L} \frac{1}{f^\tau}\right) \right. , s.t., 1 \leq l \leq L \ and \ \tau > 0 \tag{4}$$

where the τ factor characterizes the file popularity. In this study, we assume that MUs in each F-AP area request independently the lth file $\mathcal{M}_{l,1 \leq l \leq L}$. Therefore, the τ value is different for each F-AP. According to (4), $\mathcal{M}_1$ (or $\mathcal{M}_L$) has the highest (or lowest) popularity. The CS intends to rent a frequency-accessing fraction of $\mathbb{M}$ for caching to maximize the F-RAN system performance. In this study, we can denote the caching placement strategy as a two-dimensional matrix $\mathbb{I} = [0, 1]^{n \times L}$ consisting of binary entries where 1 is indicating the caching placement in a F-AP, and 0 is not. $\mathbb{I}$ is defined as

$$\mathbb{I} \triangleq \begin{bmatrix} I_1^1 & \cdots & I_1^L \\ \vdots & \ddots & \vdots \\ I_n^1 & \cdots & I_n^L \end{bmatrix} \in [0, 1]^{n \times L} \tag{5}$$

where $I_i^l = 1$ means that the file $\mathcal{M}_l$ is cached at the F - AP_i and $I_i^l = 0$ means the opposite. For the F - AP_i, the profit ($\mathfrak{R}_i^c$) gained from the local caching mechanism can be defined as follows;

$$\mathfrak{R}_i^c = \sum_{l=1}^{L} \left(g_l^i \times \mathscr{L}^i \times \mathcal{Z}_l^i \times I_i^l\right) - \sum_{l=1}^{L} \left(\mathbb{C}_l^i \times I_i^l\right), \ s.t., \ g_l^i \in \mathcal{Q}^i \tag{6}$$

where $\mathcal{Q}^i$ is the vector $\mathcal{Q}$ of F - AP_i and $\mathcal{L}^i$ is the total number of service requests on average. $\mathcal{Z}_i^l$ and $\mathbb{C}_i^l$ is the revenue and cost from the lth file caching in the F - AP_i, respectively. From the viewpoint of F - AP_i, the fraction $[I_i^1 ... I_i^L]$ of $\mathbb{I}$ ($\mathcal{Q}^i$) needs to be optimized for maximizing the $\mathfrak{R}_i^c$.

Based on the current caching placement, Service Admission Control (SAC) algorithm should be developed to make admission decisions to maximize the spectrum efficiency while maintaining a desirable overhead level. Especially, when the requested services are heavy, that is, the sum of the requested resource amount exceeds the currently available system capacity, the SAC algorithm comes into act whether to accept a new service request or not. Based on the acceptance condition, such as the current caching status and resource capacity, the SAC problem can be formulated as a joint optimization problem. In this problem, we take into account the maximization of spectrum efficiency while minimizing the fronthaul overhead.

In the proposed scheme, we set out to obtain fundamental insights into the SAC problem by means of a game theoretic approach. Therefore, the subordinated game is designed to formulate the interactions of the F-AP and MUs while investigating the system dynamics with imperfect information. To implement our subordinated game, we adopt the concept of dictator game, which is a game in experimental economics, similar to the ultimatum game, first developed by D. Kahneman et al. [23]. In the dictator game, one player, called *the proposer*, distributes his resource, and the other players, called *the responders*, simply accept the decision, which is made by *the proposer*. As one of decision theory, the dictator game is treated as an exceptional non-cooperative game or a multi-agent system game that has a partner-feature and involves a trade-off between self- and other-utility. Based on its simplicity, the dictator game can capture an essential characteristic of the repeated interaction situation [8].

In the proposed subordinated game model, each F-AP is *the proposer* and MUs are *the responders*. They interact with each other and repeatedly work together toward an appropriate F-RAN performance. To effectively make SAC decisions, *the proposer* considers the current system conditions such as the available spectrum amount, the current caching placement and fronthaul overhead status. By a sophisticated combination of these conflicting condition factors, *the proposer* attempts to approximate a temporary optimal SAC decision. The SAC decision procedure is shown in Fig. 2.

According to the SAC procedure, each F - AP_i can maintain the finest SAC solution while avoiding the heavy computational complexity or overheads. For the subordinated game, we propose a new solution concept, *Temporal Equilibrium* (*TE*). In the proposed scheme, all MUs follow compulsorily the decision of F-APs, and the outcome profile of our SAC process constitutes the *TE*, which is the current service status.

$$\begin{aligned} TE &= \overrightarrow{T\mathscr{E}}_i | (\boldsymbol{\mu}_i \cup \boldsymbol{\psi}_i) \rightarrow \left(\Theta_{j,1 \leq j \leq m}^i \in (\boldsymbol{\mu}_i \cup \boldsymbol{\psi}_i)\right) \cup \overrightarrow{T\mathscr{E}}_i \\ &= \begin{cases} \Theta_j^i \in \boldsymbol{\mu}_i, \ if \ \Theta_j^i \ is \ accepted \\ \\ \Theta_j^i \in \boldsymbol{\psi}_i, \ otherwise \end{cases} \end{aligned} \tag{7}$$

where $\overrightarrow{T\mathscr{E}}$ is the set of MUs in the F - AP_i and μ_i, ψ_i are the MUs' set of accepted or rejected by the F - AP_i, respectively. Therefore, *TE* is the status quo of dictator game.

D. The main steps of proposed F-RAN control algorithm

Define:

Θ_j^i : a new service request of MU_j^i,

$Min\_S(\Theta_j^i)$, $Min\_C(\Theta_j^i)$: the minimum spectrum and computation requirement of Θ_j^i

χ^i, ψ^i : the currently using spectrum and computation amount in the $F\text{-}AP_i$

$\mathfrak{X}^i$, $\mathfrak{M}^i$: the current and maximum fronthaul transmission rate

When the Θ_j^i is requested in the $F\text{-}AP_i$,

 If Θ_j^i request is cached in the $F\text{-}AP_i$, *// no computation offloading task*

  { **If** $\left(\chi^i + Min\_S(\Theta_j^i) \leq S_{F\text{-}AP_i}^{\mathcal{R}^i} \right)$, it is accepted

   Otherwise, it is rejected }

 Else {

  If Θ_j^i is computation offloading task,

   { **If** $\left(\left(\chi^i + Min\_S(\Theta_j^i) \leq S_{F\text{-}AP_i}^{\mathcal{R}^i} \right) \text{and} \left(\psi^i + Min\_C(\Theta_j^i) \leq S_{F\text{-}AP_i}^{\mathcal{C}^i} \right) \right)$,

    it is accepted

   Elseif $\left(\left(\chi_j^i + Min\_S(\Theta_j^i) \leq S_{F\text{-}AP_i}^{\mathcal{R}^i} \right) \text{ and } \left(\mathfrak{X}^i + Min\_S(\Theta_j^i) \leq \epsilon \times \mathfrak{M}^i \right) \right)$

    it is accepted,

   Otherwise, it is rejected }

  Elseif $\left(\left(\chi_j^i + Min\_S(\Theta_j^i) \leq S_{F\text{-}AP_i}^{\mathcal{R}^i} \right) \text{ and } \left(\mathfrak{X}^i + Min\_S(\Theta_j^i) \leq \epsilon \times \mathfrak{M}^i \right) \right)$

    it is accepted,

  Otherwise, it is rejected

 }

Fig. 2 Service Admission Control procedure

For 5G wireless communications, the F-RAN architecture is a promising paradigm to provide high spectrum efficiency and improved QoS. The core idea of F-RAN is to take full advantages of distributed edge processing, cooperative spectrum allocation, and reduced communication latency [24–26]. However, it is questionable whether existing F-RAN control schemes are adequate in dynamically changing F-RAN environments. In this study, we have studied the joint design of spectrum allocation and SAC decision in an F-RAN architecture. Focusing on the practical assumption, we develop a new embedded game model while investigating the benefits and challenges of F-RAN control mechanisms. In our embedded game approach, the superordinated game for spectrum allocations and the subordinated game for SAC decisions are interlocked and serially correlated. The subordinated game depends on the outcome of the superordinated game, and the result of subordinated games is the input back to the superordinated game process. Structurally, the multiple subordinated games are nested in the superordinated game, and they are linked based on the step-by-step interactive feedback process. It may be the only realistic approach to solve complex and dynamically changing F-RAN control problems. Usually, the traditional optimal and centric algorithms have exponential time complexity. However, the proposed distributed control method has only polynomial time complexity. The main steps of the proposed F-RAN control algorithm are given next (see Fig. 3).

Step 1: At the initial time, the spectrum resource allocation $S_{cs}^{\mathcal{R}} = \{\delta_1, \delta_2, \dots \delta_n\}$ and trustworthiness ($\mathcal{T}$) for F-APs are equally distributed. This starting guess guarantees that each F-AP enjoys the same benefit at the beginning of the game.

Step 2: Control parameters are given from the simulation scenario (refer to Table 2). To fairly compare with the existing schemes, the system parameters are carefully selected in our simulation model.

Step 3: By taking into account the current F-RAN situations, our superordinated and subordinated games are executed in parallel.
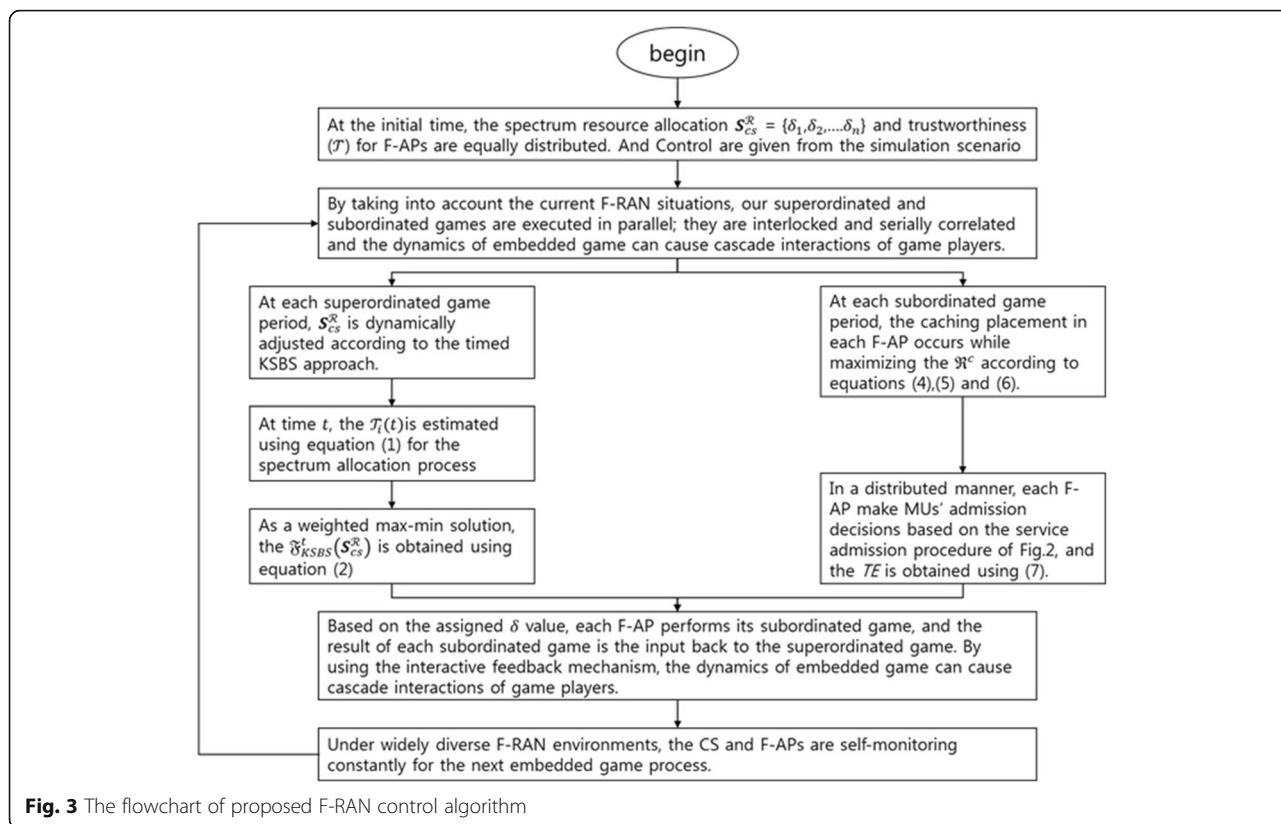
Fig. 3 The flowchart of proposed F-RAN control algorithm

Table 2 System parameters used in the simulation experiments

Application type	Computation offloading	Computation requirement	Minimum spectrum requirement	Maximum spectrum requirement
I	Y	300 MHz/s	128 kbps	128 kbps
II	N	N/A	256 kbps	768 kbps
III	Y	600 MHz/s	384 kbps	640 kbps
IV	N	N/A	512 kbps	1.28 Mbps

Parameter	Value	Description
n	10	The number of F-APs
$\mathcal{R}_{CS}$	200 Mbps	The total spectrum resources of CS
$\mathcal{C}$	5 GHz	The F-AP's computation capacity
φ	0.2	A factor to specify the impact of recent experience
Δt	1 s	The time interval to monitor the F-RAN system
$\mathcal{Z}$	5 / one bps	The revenue from the caching per one bps
$\mathfrak{C}$	1 / one bps	The cost from the caching per one bps
τ	[0.1–0.9]	A factor to characterize the file popularity: randomly selected for F-AP
L	10	The popular multimedia files in the CS for caching
$\mathfrak{M}$	30 Mbps	The maximum fronthaul transmission rate
	0.95	A control factor to consider the fronthaul congestion

Step 4: The trustworthiness ($\mathcal{T}$) for each F-AP is modified periodically by using (1).

Step 5: At each superordinated game period, $S_{cs}^{\mathscr{R}} = \{\delta_1, \delta_2, ... \delta_n\}$ is dynamically adjusted according to the timed KSBS manner. According to (2), the $\mathfrak{F}_{KSBS}^{t}(S_{cs}^{\mathscr{R}})$ is obtained, and each δ value is decided for the next game period.

Step 5: At each subordinated game period, the caching placement in each F-AP occurs while maximizing the $\mathfrak{R}^c$ according to Eqs. 4, 5 and 6.

Step 7: In a distributed manner, each F-AP makes MUs' admission decisions based on the service admission procedure of Fig. 2, and the *TE* is obtained using (7).

Step 8: The superordinated and subordinated games are interlocked and serially correlated. Based on the assigned δ value, each F-AP performs its subordinated game, and the result of each subordinated game is the input back to the superordinated game.

Step 9: Based on the interactive feedback mechanism, the dynamics of embedded game can cause cascade interactions of game players and players can make their decisions to quickly find the most profitable solution.

Step 10: Under widely diverse F-RAN environments, the CS and F-APs are self-monitoring constantly for the next embedded game process; proceed to step 3.

4 Performance evaluation

In this section, we compare the performance of our F-RAN control scheme with other existing schemes [3, 6, 17] and can confirm the performance superiority of the proposed approach by using a simulation model. To fairly compare the system performance, the assumptions and detailed system scenario are outlined as follows.

- The simulated system consists of one CS, 10 F-APs and multiple MUs. The number of MUs (m) for each F-AP is generated based on the process for new service requests.
- The process for new service requests is Poisson with rate λ (services/s), and the range of offered service load was varied from 0 to 3.
- There are four different service applications. They are randomly generated from MUs, and some of them are computation offloading tasks.
- The durations of service applications are exponentially distributed.
- The total spectrum resources of CS ($\mathscr{R}_{CS}$) is 200 Mbps.
- For each F-AP, the computation capacity ($\mathcal{C}$) is 5 GHz, and the fronthaul link capacity is 30 Mbps.

- The cache size in each F-AP is the same as the file set $\mathbb{M}$ in the CS.
- System performance measures obtained on the basis of 100 simulation runs are plotted as functions of the service generation rate.
- For simplicity, we assume the absence of physical obstacles in the experiments.

Performance measures obtained through simulation are the normalized F-RAN throughput, spectrum efficiency, and fronthaul transmission delay. To facilitate the development and implementation of our simulator, Table 2 lists the system control parameters.

Figure 4 shows the performance comparison of each scheme in terms of the normalized system throughput. It is estimated as the total data transmission in the F-RAN system. From the viewpoint of system operator, it is a key factor in the F-RAN management. It can be seen from Fig. 4 that the throughput of all the schemes increases as the service request rate increases, and we can confirm the performance superiority of our scheme. The proposed scheme's gain on performance can be obtained through (i) the effective coordination paradigm by employing an embedded game model, and (ii) the jointed design of spectrum allocation and SAC decision algorithms to obtain synergistic and complementary features. Therefore, our scheme can get a better performance than other existing schemes, which were designed as one-sided protocols and did not consider the feasibility to respond the current F-RAN system conditions.

In Fig. 5, we plot the spectrum efficiency. It means a bandwidth usage ratio of F-RAN system. In general, as the service request rate increases, the spectrum efficiency also increases. This is intuitively correct. In our embedded game approach, all system agents adaptively interact with each other and decide their strategy based on the current F-RAN system conditions. Therefore, we allocate the spectrum resource based on the timed KSBS approach, and can maintain a higher spectrum efficiency. Figure 5 clearly indicates that the proposed scheme can effectively handle the resource allocation problem than the existing schemes in [3, 6, 17] from low to heavy service load distributions.
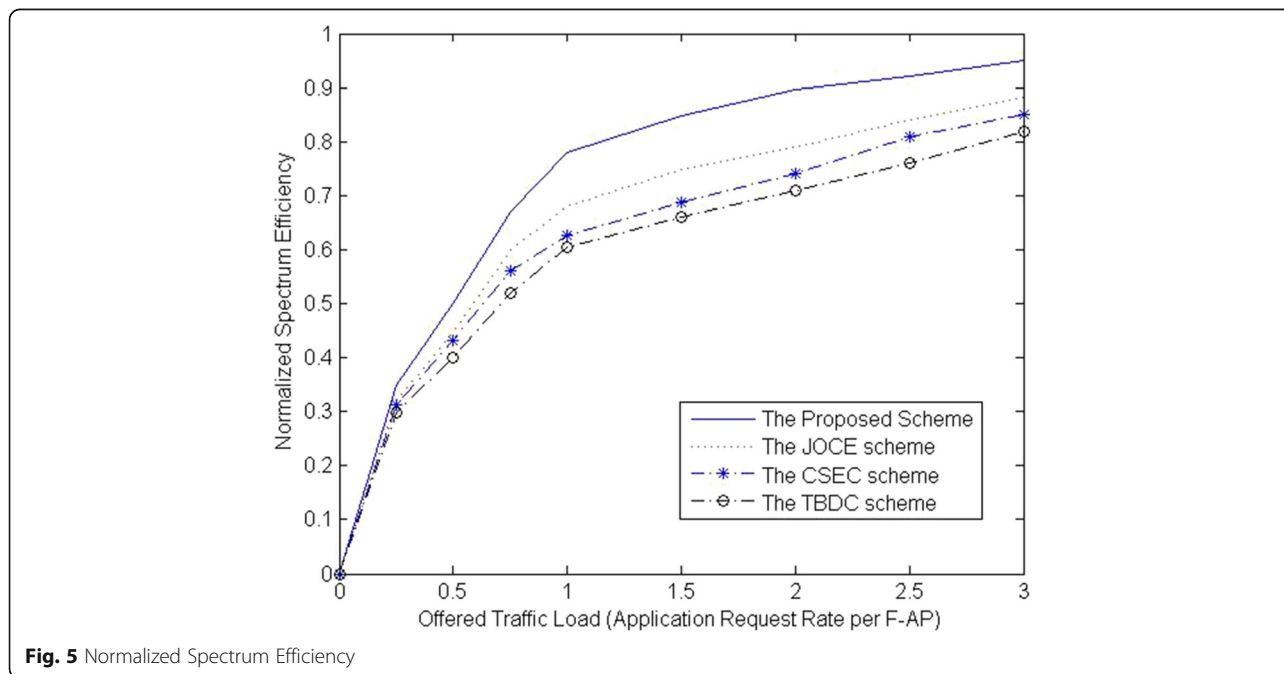
Figure 6 shows the fronthaul transmission delay curves for the data communication implemented with four different choices of each schemes. It is estimated as the normalized time delay between the CS and its corresponding MU. In order to quantify the F-RAN's QoS performance, it is one of important metrics. The result shows that the proposed scheme with an adaptive SAC mechanism can achieve a significantly lower transmission delay. The simulation results shown in Figs. 4, 5 and 6 demonstrate the performance comparison of the proposed scheme and other existing schemes [3, 6, 17],

Fig. 4 Normalized F-RAN Throughput
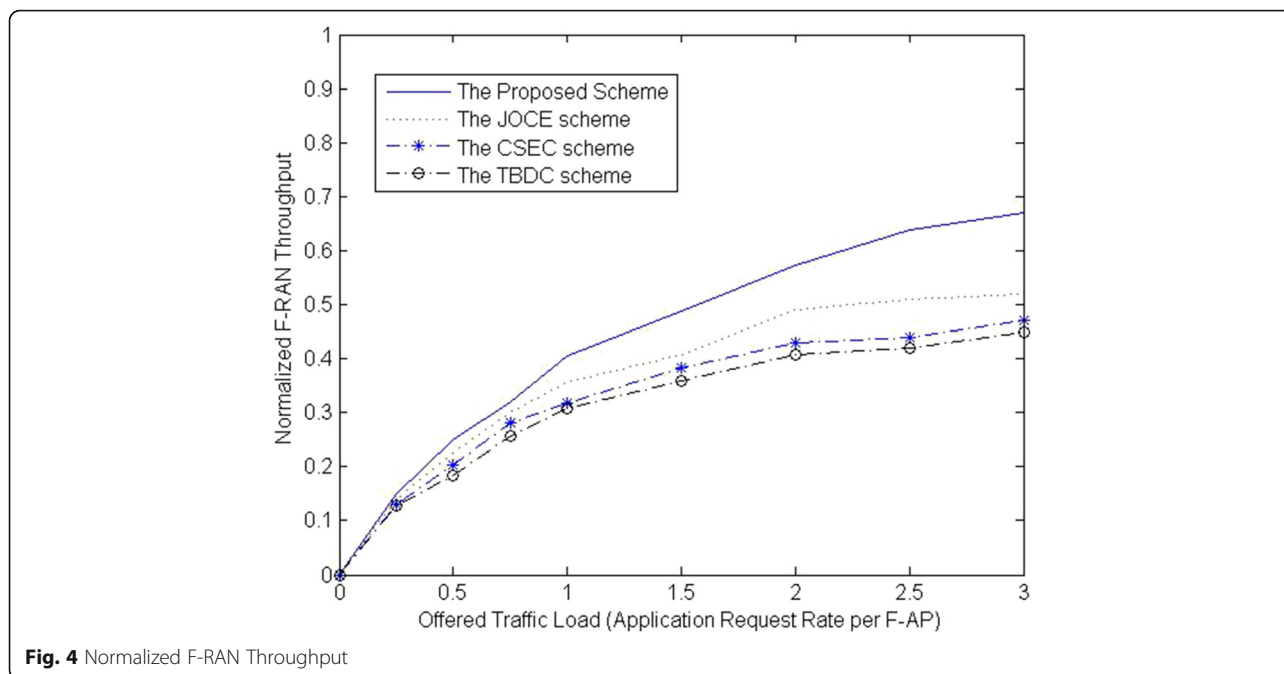
and verify that the proposed embedded game approach can strike the appropriate performance balance between system throughput, spectrum efficiency, and transmission delay; the *Joint Optimization of Cloud and Edge* (*JOCE*) scheme [3], the *Cloud Structure with Edge Caching* (*CSEC*) scheme [6], and the *Traffic Balancing and Dynamic Clustering* (*TBDC*) scheme [17] cannot offer such an attractive performance balance.

5 Conclusions

As a promising paradigm for the 5G communication system, the F-RAN has been proposed as an advanced socially aware wireless networking architecture to provide the higher spectral efficiency while maximizing the system performance. In this study, we have studied joint design of cloud and edge processing in the F-RAN system to solve the resource allocation

Fig. 5 Normalized Spectrum Efficiency
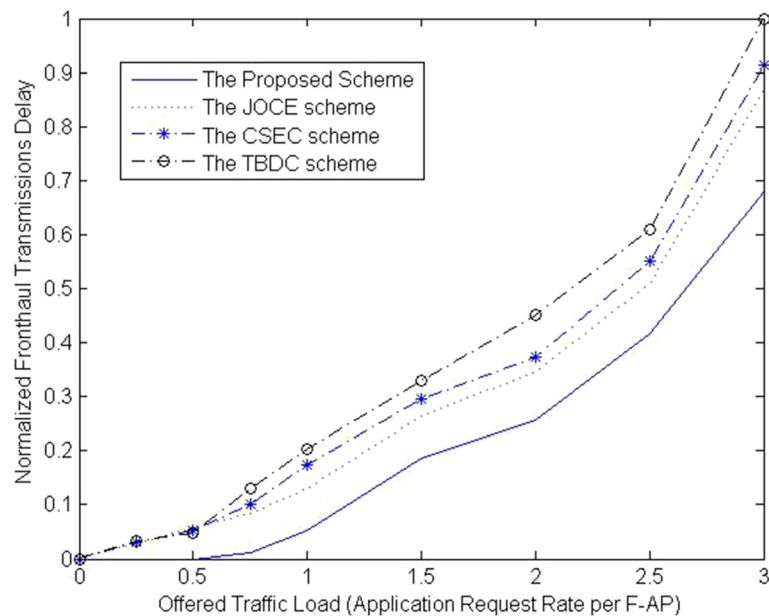
Fig. 6 Normalized Fronthaul Transmission Delay

and SAC problems. Based on the newly developed embedded game model, we have explored the feasibility of F-RAN control decision process and the practicality for the real-world implementation. In our embedded game structure, the SAC algorithm is nested in the spectrum allocation algorithm to effectively control the conflict problem of F-RAN system agents. Based on the interactive feedback mechanism, the proposed scheme has the potential to handle multiple targets without using more complex multi-target tracking algorithm. The extensive simulation result is very encouraging, showing that our embedded game-based approach provides a more effective way to control the F-RAN system than the other existing schemes. Open issues for the further research are the designs and validations of the original F-RAN systems for big data mining, cognitive radio, software-defined network, and network security problems. The progress of trial tests and test bed development of F-RANs can be anticipated to be promoted in the future, which makes F-RANs' commercial rollout as early as possible.

Author's contribution
SK: is a sole author of this work and ES (i.e., participated in the design of the study and performed the statistical analysis).

Competing interests
The author, Sungwook Kim, declares that there is no competing interests regarding the publication of this paper.

6 Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. S-C Hung, H Hsu, S-Y Lien, K-C Chen, Architecture Harmonization Between Cloud Radio Access Networks and Fog Networks. IEEE Access **3**, 3019–3034 (2015)
2. R Tandon, O Simeone, Harnessing cloud and edge synergies: toward an information theory of fog radio access networks. IEEE Commun. Mag. **54**(8), 44–50 (2016)
3. S-H Park, O Simeone, S Shamai, *Joint optimization of cloud and edge processing for fog radio access networks* (IEEE ISIT, 2016), pp. 315–319
4. AV Dastjerdi, R Buyya, Fog Computing: Helping the Internet of Things Realize Its Potential. Computer **49**(8), 112–116 (2016)
5. P Borylo, A Lason, J Rzasa, A Szymanski, A Jajszczyk, *Energy-aware fog and cloud interplay supported by wide area software defined networking* (IEEE ICC, 2016), pp. 1–7
6. R Tandon, O Simeone, *Cloud-aided wireless networks with edge caching: Fundamental latency trade-offs in fog Radio Access Networks* (IEEE ISIT, 2016), pp. 2029-2033
7. Z Hu, Z Zheng, T Wang, L Song, X Li, Game theoretic approaches for wireless proactive caching. IEEE Commun. Mag. **54**(8), 37–43 (2016)
8. S Kim, *Game Theory Applications in Network Design* (IGI Global, Hershey, 2014)
9. K Sundaresan, MY Arslan, S Singh, S Rangarajan, SV Krishnamurthy, FluidNet: A Flexible Cloud-Based Radio Access Network for Small Cells. IEEE/ACM Trans. Networking **24**(2), 915–928 (2016)
10. J Wu, Z Zhang, H Yu, Y Wen, Cloud radio access network (C-RAN): a primer. IEEE Netw. **29**(1), 35–41 (2015)
11. Q-T Vien, N Ogbonna, HX Nguyen, R Trestian, P Shah, in *Non-Orthogonal Multiple Access for Wireless Downlink in Cloud Radio Access Networks*, Proceedings of European Wireless (2015), pp. 1-6

12. Q-T Vien, TA Le, B Barn, CV Phan, Optimising energy efficiency of non-orthogonal multiple access for wireless backhaul in heterogeneous cloud radio access network. IET Commun. **10**(18), 2516–2524 (2016)

13. HQ Tran, PQ Truong, CV Phan, Q-T Vien, *On the energy efficiency of NOMA for wireless backhaul in multi-tier heterogeneous CRAN* (SigTelCom, 2017), pp. 229–234

14. M Peng, Y Li, J Jiang, J Li, C Wang, Heterogeneous cloud radio access networks: a new perspective for enhancing spectral and energy efficiencies. IEEE Wirel. Commun. **21**(6), 126–135 (2014)

15. A Sengupta, R Tandon, O Simeone, *Fog-Aided Wireless Networks for Content Delivery: Fundamental Latency Trade-Offs*, (2015) [Online]. Available: https://arxiv.org/abs/1605.01690. Accessed 16 Apr 2017.

16. SM Azimi, O Simeone, A Sengupta, R Tandon, *Online Edge Caching in Fog-Aided Wireless Network*, (2017) [Online]. Available: https://arxiv.org/abs/1701.06188. Accessed 16 Apr 2017.

17. D Chen, S Schedler, V Kuehn, *Backhaul traffic balancing and dynamic content-centric clustering for the downlink of Fog Radio Access Network* (IEEE SPAWC, 2016), pp. 1–5

18. M Peng, S Yan, K Zhang, C Wang, Fog-computing-based radio access networks: issues and challenges. IEEE Netw. **30**(4), 46–53 (2016)

19. H Qiao, J Rozenblit, in *Ferenc Szidarovszky and Lizhi Yang, Multi-Agent Learning Model with Bargaining*, Proceedings of the 2006 Winter Simulation Conference, pp. 934-940, 2006

20. F Bao, I-R Chen, *Trust management for the internet of things and its application to service composition* (IEEE WoWMoM, 2012), pp. 1–6

21. K Sungwook, News-vendor game-based resource allocation scheme for next-generation C-RAN systems. EURASIP J. Wirel. Commun. Netw. **2016**(1), 1–11 (2016)

22. J Li, J Sun, Y Qian, F Shu, M Xiao, W Xiang, A Commercial Video-Caching System for Small-Cell Cellular Networks using Game Theory. IEEE Access **4**, 7519–7531 (2016)

23. D Kahneman, JL Knetsch, RH Thaler, Fairness and the assumptions of economics. J. Bus. **59**(4), 285–300 (1986)

24. W Zhu, C Lee, A New Approach to Web Data Mining Based on Cloud Computing. JCSE **8**(4), 181–186 (2014)

25. Y Liu, Y Sun, J Ryoo, S Rizvi, AV Vasilakos, A Survey of Security and Privacy Challenges in Cloud Computing: Solutions and Future Directions. JCSE **9**(3), 119–133 (2015)

26. K Lee, I Shin, User Mobility Model Based Computation Offloading Decision for Mobile Cloud. JCSE **9**(3), 155–162 (2015)