Dynamic Service Selection and Bandwidth Allocation in IEEE 802.16m Mobile Relay Networks

Kun Zhu, Dusit Niyato, and Ping Wang

Abstract

Cooperative relay network will be supported in IEEE 802.16m to improve the coverage and performance of mobile broadband wireless access service. In this paper, we jointly consider the problem of dynamic service selection and bandwidth allocation in IEEE 802.16m mobile relay networks. Specifically, the advanced mobile stations (AMSs) perform the selection of advanced base station (ABS) and transmission mode (i.e., direct transmission or relay-cooperation transmission) for a better service quality. The ABSs allocate different bandwidth for different transmission modes to maintain the desired queue level at base stations and user distribution for satisfying performance requirements. This problem is challenging when the strategies of both ABSs and AMSs influence each other and the decisions are made dynamically. To address this problem, a two-level dynamic game framework based on an evolutionary game and a differential game is developed. Since the mobile stations can adapt their strategies according to the received service quality, the dynamic service selection is modeled as an evolutionary game at the lower level. At the upper level, a differential game is formulated for a dynamic bandwidth allocation of base stations and a closed-loop Nash equilibrium is obtained as the solution. Viewing the fluctuation of traffic flow rate as disturbance, the robust bandwidth allocation strategy design is performed. Both stochastic optimal control and $H_\infty$ optimal control approaches are adopted for average performance and worst-case performance design, respectively. The developed framework relies on minimum information which makes the system flexible and simple for implementation. Numerical results show the effectiveness of dynamic game equilibrium bandwidth allocation strategies and also the advantages in terms of performance and convergence speed.

Keywords – Relay cooperative communication, Dynamic bandwidth allocation, Replicator dynamics, Differential game, Closed-loop Nash equilibrium.

K. Zhu, D. Niyato, and P. Wang are with the School of Computer Engineering, Nanyang Technological University (NTU), Singapore. D. Niyato is the corresponding author (email: dniyato@ntu.edu.sg).
I. Introduction

IEEE 802.16m is proposed as an enhancement for the existing mobile WiMAX systems to meet the requirements of next-generation mobile broadband communication in the framework of IMT-Advanced [1]. Relay transmission as a cost-effective approach for extending the coverage area and enhancing the overall throughput is specified in IEEE 802.16m standard [2], [3]. Relay-cooperation transmission can be seen as a kind of cooperative communications, in which a relay station helps to forward data between mobile stations and base stations. In IEEE 802.16m, the advanced base station (ABS) frame (Fig. 1) is divided into access zone and relay zone for supporting both advanced base station - advanced relay station (ABS-ARS) communications and advanced base station - advanced mobile station (ABS-AMS) communications, and the bandwidth of ABS in IEEE 802.16m for both communications has to be allocated accordingly. Due to the asymmetry of downlink (DL) and uplink (UL) traffic load, the duration of the access and relay zones could be different.

The bandwidth allocation for ABS-ARS and ABS-AMS communications is challenging due to the facts that the bandwidth allocation strategies of ABSs and the service selection strategies of AMSs influence each other and the decisions are made dynamically under competition. In particular, AMS can choose ABSs and services modes (i.e., direct transmission or relay-cooperation transmission) dynamically while the ABSs can competitively allocate the bandwidth in a dynamic manner such that their individual target system performance can be achieved. The former (defined as the service selection problem) should be performed by taking the time-varying user performance into account so that the utility of AMS is maximized. Meanwhile the latter (defined as the bandwidth allocation problem) should be done by considering the dynamic behavior of AMSs to maintain the desired queue level at the base stations and user distribution for satisfying system performance requirements. The service selection and bandwidth allocation are intertwined in such a way that the bandwidth will determine the preference of the users in selecting the services. Also, the user service selection affects the performance of the network. The game theory can be applied to analyze this situation.

To jointly address the issues of dynamic service selection and dynamic bandwidth allocation in IEEE 802.16m mobile relay networks, the novel hierarchical (i.e., two-level) dynamic game framework based on evolutionary game [4] and differential game [5] with inherent dynamic nature is developed to obtain

\(^{1}\)In the rest of the paper, the terminologies AMS, ABS, and ARS in IEEE 802.16m are used interchangeably with mobile station (MS), base station (BS), and relay station (RS) in IEEE 802.16j, respectively.
the equilibrium strategies for both AMSs and ABSs. At the lower level, the dynamic competition for resource access among AMSs is modeled as an evolutionary game since the AMSs can adapt their service selection strategies according to the received service quality (i.e., performance). The service selection strategy adaptation of this evolutionary game is subject to the control of ABSs in terms of allocated bandwidth, and is modeled by replicator dynamics [6]. Evolutionary equilibrium is considered to be the solution of this lower level evolutionary game. At the upper level, the ABSs observe the service selection distribution of the AMSs and the queue states at the ABSs which jointly describe the system state, and then decide the optimal bandwidth allocation strategies dynamically. This dynamic bandwidth allocation is formulated as a differential game and a closed-loop Nash equilibrium is considered to be the solution. Since the rate of traffic input flows can fluctuate over time, the robust bandwidth allocation strategy design is conducted by viewing the rate fluctuation as a disturbance. In this case, both stochastic optimal control approach and $H_{\infty}$ optimal control approach are applied for deriving the robust optimal strategies depending on different disturbance characteristics. Specifically, the stochastic optimal control is adopted for the average performance design when the statistics of the disturbance appear as a white noise Gaussian process, while the $H_{\infty}$ optimal control is adopted for the worst-case performance design under the general finite energy disturbances.

The main motivation of this two-level dynamic game framework stems from the consideration of the dynamic decision making for both AMSs and ABSs. The novelty is the use of differential game for the upper level dynamic competition and robust bandwidth allocation strategy design. The proposed dynamic game framework has several favorable properties. First, the inherent dynamic nature makes this framework applicable to the competitive dynamic decision making problems with hierarchical structure which are common in wireless networks. Second, fairness can be achieved by using the replicator dynamics for the evolutionary game in which all AMSs receive the same individual utility. Third, comparing with static strategies (e.g., the Nash equilibrium of a static non-cooperative bandwidth allocation game), the higher efficiency (i.e., better performance) and the faster convergence speed to steady states are the main advantages of using the dynamic bandwidth allocation strategies for ABSs.

The rest of this paper is organized as follows: Section II provides a survey on the related work. Section III presents the system model. The evolutionary game formulation for a strategy adaption of the AMSs is presented in Section IV. The competitive bandwidth allocation considering the dynamic service selection is formulated as a differential game in Section V. The derivation of robust optimal allocation strategies based
on the stochastic optimal control and $H_\infty$ optimal control is also presented in this section. Section VI presents the numerical results. The conclusions are drawn in Section VII.

II. RELATED WORK

A. IEEE 802.16j/m Mobile Relay Networks

The performance of relay-based cooperative communications highly relates to the issues such as frame design, networking planning, relay selection, and bandwidth allocation. A few works have been done on these issues in IEEE 802.16j/m mobile relay networks. [7] proposed a frame structure for multihop communications in IEEE 802.16j with backward compatibility with the legacy IEEE 802.16 standards. [8], [9] jointly considered the problems of network planning and bandwidth allocation in IEEE 802.16j relay networks. [10], [11] addressed the relay selection issue, i.e., when to cooperate and whom to cooperate with. In [12], the problem of joint relay selection and link adaption in IEEE 802.16j/m was addressed. In [13], the selection of an optimal relay station for a vehicular subscriber station was studied. The problems of bandwidth allocation and frame partitioning in IEEE 802.16j were studied in [14], [15]. However, none of the works jointly considered the problems of dynamic service selection of advanced mobile stations and dynamic bandwidth allocation of advanced base stations in IEEE 802.16m mobile relay networks. The key difference between our work and previous approaches is the dynamic robust optimal\(^2\) decision making for base stations and dynamic resource competition among mobile stations.

B. Evolutionary and Differential Games in Wireless Networks

Evolutionary game theory developed as the mathematical tool to study the population dynamics has been widely applied in wireless networks. [16]–[20] addressed different issues (e.g., routing, access network selection, energy management, and congestion control) in wireless networks using evolutionary game theory. Differential game belongs to a subclass of dynamic games and it is used to study the dynamic decision making problems in competitive systems that evolve over continuous time. A linear state differential game was formulated to address the problem of dynamic bandwidth allocation in heterogeneous wireless networks in our previous work [21]. Optimal control and $H_\infty$ robust control as the approaches for deriving the Nash equilibrium of differential game have wide applications (e.g., bandwidth allocation, rate control, routing, and power control) in wireless networks [22]–[25]. However, to the best of our

\(^2\)Throughout the paper, “optimal” means best response to the strategies of competitors in the competitive environment.
knowledge, evolutionary and differential games have not been considered jointly and applied to solve radio resource management problem in wireless networks.

III. System Model

A. IEEE 802.16m Relay Network

We consider the following system of IEEE 802.16m relay network to address the problem of dynamic service selection jointly with bandwidth allocation. There is a service area $a$ in an IEEE 802.16m mobile relay network with $M$ competitive access providers and $N$ advanced mobile stations (AMSs) as shown in Fig. 2. Without loss of generality, each access provider deploys one advanced base station (ABS), and hence, there are totally $M$ ABSs providing service in this service area. Access provider also deploys advanced relay stations (ARSs) in this area to compete for providing an access service to AMSs. The current IEEE 802.16m specification only includes a subset of relay categories of IEEE 802.16j and only fixed ARS is considered [3]. Also, the ARSs in IEEE 802.16m are operated in a non-transparent mode and only two-hop communication is considered for relay-cooperation transmission. We assume that the ARSs associated with the ABS of the same provider are deployed densely and evenly in the area to make the access indifferent to the roaming AMSs.\(^3\) Two transmission modes (i.e., direct transmission denoted by ABS-AMS and relay-cooperation transmission denoted by ABS-ARS-AMS) can be used for AMSs. In this case, an AMS can choose to access the service either from an ABS directly or with the cooperation of an ARS. Also, the AMSs can select and churn to different access providers (i.e., ABSs) freely and dynamically. In the following, the two transmission modes are regarded as two service types provided by the ABS from the AMS perspective. Therefore, there are in total $2M$ choices for service selection of an AMS.

Consider downlink transmission\(^4\), let $C_i$ denote the total downlink bandwidth of ABS $i$. In IEEE 802.16m, the downlink (DL) subframe is divided into zones (i.e., access zone and relay zone), and the total downlink bandwidth of the ABS will be partitioned to support both ABS-AMS and ABS-ARS-AMS communications (Fig. 2). Let $\mu_{ij}(t)$ denote the allocated bandwidth of ABS $i$ at time $t$ for different transmission mode where $i \in \{1, 2, \ldots, M\}$, $j = 1$ for direct and $j = 2$ for relay-cooperation transmissions. For the resource allocation among AMSs, we consider an un-weighted round-robin scheme

\(^3\)The issue of relay selection and deployment is out of the scope of this paper.
\(^4\)The model can be extended for uplink direction.
where all AMSs choosing the same service type of an ABS share the available bandwidth equally. Therefore, the bandwidth of an AMS allocated from ABS $i$ is denoted by $b_{ij} = \mu_{ij}(t)/n_{ij}(t)$ where $n_{ij}(t)$ represents the number of AMSs choosing direct ($j = 1$) and relay-cooperation transmission ($j = 2$) of ABS $i$ at time $t$.

The selection of ABS (i.e., access provider) as well as transmission mode of an AMS depends on the received performance which is dependent of the transmission rate and channel quality. Let $\tau_{ij}$ and $k_{ij}$ denote the average downlink transmission rate and average spectral efficiency (in bits/s/Hz) for ABS $i$ with transmission mode $j$, respectively. The spectral efficiency can be obtained based on the channel estimation. The average downlink transmission rate of an AMS in area $a$ with direct communication with ABS $i$ is obtained as $\tau_{i1} = b_{i1}k_{i1}$, where $b_{i1}$ is the allocated bandwidth from ABS $i$ and $k_{i1}$ is the average spectral efficiency for direct communication. The average downlink transmission rate for an AMS with relay-cooperation communication with ABS $i$ is obtained as $\tau_{i2} = b_{i2}\min\{k_{i2}, k_{im}\}$, where $b_{i2}$ is the allocated bandwidth from ABS $i$ for relay-cooperation transmission, $k_{i2}$ and $k_{im}$ (subscript $im$ stands for ARS-AMS) are the average spectral efficiency for ABS-ARS and ARS-AMS communication from ABS $i$, respectively. Basically, the rate of relay-cooperation transmission is the minimum rate between ABS-ARS and ARS-AMS. For the performance satisfaction level of an AMS, utility function is used. Specifically, we consider a linear function $u_{ij} = \alpha b_{ij}k_{ij}(1 - p_{ij})$, where $\alpha$ is an adjustable constant to provide flexibility to fit the function with empirical data for different types of user applications, and $p_{ij}$ is the average packet error rate (PER) for abstracting the channel quality. The estimation of PER can be found in [26].

### B. Queue Dynamics

For the downlink data transmission, we assume that there is one queue associated with each service type of the ABS at a link layer (Fig. 2). Denote $q_{ij}(t)$ the queue length of aggregated traffic flow of service type $j$ at ABS $i$ at time $t$, where $j \in \{1, 2\}$ for direct and relay-cooperation transmissions, respectively. The vector of the queue states of all ABSs is then given by $q = [q_{11}(t), q_{12}(t), \ldots, q_{M2}(t)]^T$. Similar to [24], a linear fluid model\(^5\) is considered for modeling the queue dynamics, and the queue length evolves according to

$$
\dot{q}_{ij}(t) = n_{ij}(t)r(t) - (1 - p_{ij})\mu_{ij}(t)k_{ij},
$$

\(^5\)This model is a continuous time version of the models used in [22], [27].
for all $i \in \{1, 2, \ldots, M\}, j \in \{1, 2\}$, where $\dot{q}_{ij}(t)$ is the derivative of $q_{ij}(t)$ with respect to $t$, $r(t)$ is the input traffic flow rate to the subscribed AMS, $n_{ij}(t)r(t)$ is the aggregate downlink flow rate at ABS $i$, and $\mu_{ij}(t)k_{ij}$ is the queue depletion rate (assigned bandwidth for draining the queue). The initial state $q_{ij}(0)$ represents the initial size of backlogged data of the queue. The input rate $r(t)$ can be fluctuated over time depending on the source behavior (e.g., variable bit rate (VBR) traffic) and can be viewed as the disturbance to the system. Specifically, we denote

$$r(t) = r_n + \xi(t), \quad (2)$$

where $r_n$ is a nominal value of the input rate and $\xi(t)$ is a disturbance which can be either stochastic (e.g., white noise Gaussian process) or deterministic (e.g., impulse traffic load). This disturbance can occur due to the randomness of the packet arrival from the applications. It has been shown that the aggregated traffic (e.g., from users and applications) can be accurately modeled as the Gaussian process [28], [29]. From (1), we can see that both the bandwidth allocation strategies of ABSs and the service selection decisions of AMSs affect the queue states evolution. One objective of the dynamic bandwidth allocation is to maintain the queue length at the target level for satisfying system performance requirements.

C. Hierarchical Dynamic Game Model

To address the problems of dynamic service selection and downlink bandwidth allocation in IEEE 802.16m mobile relay networks, the hierarchical dynamic game framework based on an evolutionary game and a differential game is proposed. The framework is composed of two levels described as follows:

- **Lower-level evolutionary game**: An evolutionary game is used to model the competition for wireless resources among users (e.g., AMSs). The players are the AMSs in the service area (i.e., population). The strategy is to select the transmission mode (i.e., ABS-AMS or ABS-ARS-AMS) and the ABS (i.e., access provider). The details of this lower-level evolutionary game will be provided in Section IV. Note that the evolutionary game is a suitable tool for this case due to the following reasons. First, the number of AMSs involved in this competition is commonly large which is difficult to handle in traditional non-cooperative game theory. Second, the resource competition is naturally dynamic due to the characteristics of wireless networks and traffic source. Third, the AMSs in the lower level usually lack of complete information and computation capability.

6For simplicity, we assume one downlink flow for each AMS and the source rate characteristics are the same for all flows.
• **Upper-level differential game:** Based on the service selection (i.e., lower-level evolutionary game), the upper-level differential game is applied to model the competition among ABSs (which are the players in the upper-level differential game). The strategy is the bandwidth allocation of ABS for direct transmission and relay-cooperation transmission. The objective of ABS is to guarantee the performance requirements (i.e., described by a cost function which will be discussed in Section V in detail). Due to the dynamic optimal decision making process in competitive environment, the differential game is a suitable tool for this case.

In the proposed two-level framework, the service selection distribution (i.e., population state) of the lower-level evolutionary game and the queue states of ABSs jointly describe the state of the upper-level differential game. This state describes the interrelationships between the two levels. On one hand, the evolution of population state and queue dynamics is governed by the control strategies of the upper-level players (i.e., ABSs). On the other hand, in the upper-level differential game, the derivation of optimal strategies of players (i.e., AMSs) needs to take the population state evolution into account as the constraints.

**IV. Service Selection Evolutionary Game of Advanced Mobile Stations**

We consider a large population of advanced mobile stations (AMSs) in service area \( a \) competing for the downlink bandwidth resources by selecting different service types (i.e., transmission modes) and different advanced base stations (ABSs). The objective of this selection is to maximize the individual satisfaction (i.e., utility) in terms of performance. In this section, first the evolutionary game formulation and service selection dynamics of AMSs are presented. The important properties (i.e., existence and uniqueness of the population state evolution) are discussed. Then, the equilibrium and evolutionary stability are analyzed.

**A. Service Selection Dynamics**

The players of a lower-level evolutionary game are the \( N \) AMSs in area \( a \). The strategy of each player is the selection of ABS and the service type. Let the normalized variable \( x_{ij}(t) = n_{ij}(t)/N \) denote the proportion of AMSs in area \( a \) choosing service type \( j \) from ABS \( i \) at time \( t \), where \( i \in \{1, 2, \ldots, M\} \) and \( j \in \{1, 2\} \). The population state is denoted by a vector \( \mathbf{x}(t) = [x_{11}(t), x_{12}(t), \ldots, x_{M2}(t)]^T \in \mathbb{X} \), where \( x_{ij}(t) \in [0, 1], \sum_{i=1}^{M} \sum_{j=1}^{2} x_{ij}(t) = 1 \) for all \( t \geq 0 \), and \( \mathbb{X} \) is the state space. The payoff (i.e., fitness) of AMS as a player is the utility that quantifies the performance satisfaction level according to the transmission rate and channel quality. Denote \( \mathbf{\mu}(t) = [\mu_{11}(t), \mu_{12}(t), \ldots, \mu_{ij}(t), \ldots, \mu_{M1}(t), \mu_{M2}(t)]^T \) the
bandwidth allocation control vector. The expected payoff of an AMS selecting service type \( j \) of ABS \( i \) can be defined as \( \pi_{ij}(x, \mu) = u(\tau_{ij}(x, \mu)) \).\(^7\) The weighted average payoff of the population can be derived as \( \pi(x, \mu) = \sum_{i=1}^{M} \sum_{j=1}^{2} c_{ij} x_{ij}(t) \pi_{ij}(x, \mu) \), where \( c_{ij} \) is the weighting factor inverse proportional to the cost (e.g., power) for obtaining current service quality. For simplicity, we consider \( c_{ij} = \beta/k_{ij}(1-p_{ij}) \), where \( \beta \) is a constant coefficient.

To model the strategy adaption process, replicator dynamics represented by a set of differential equations are used. In the replicator dynamics, the variation rate of the population share of a particular strategy is proportional to the difference between the payoff of that strategy and the average payoff of the population.

The replicator dynamics for service selection are defined as follows:

\[
\dot{x}_{ij}(t) = f_{ij}(x, \mu) = \delta x_{ij}(t) \left( \pi_{ij}(x, \mu) - \overline{\pi}(x, \mu) \right),
\]

with initial condition \( x(0) = x_0 \in \mathbb{X} \), for all \( i \in \{1, 2, \ldots, M\}, j \in \{1, 2\} \), where \( \dot{x}_{ij}(t) \) is the derivative of \( x_{ij}(t) \) with respect to \( t \), and \( \delta \) is the learning rate of the population which controls the speed of change in service selection.

\textbf{Remark 1:} The constraints \( \sum_{i=1}^{M} \sum_{j=1}^{2} x_{ij}(t) = 1 \) and \( x_{ij}(t) \in [0, 1] \) can be guaranteed for all \( t \in [0, \infty) \) under initial condition \( \sum_{i=1}^{M} \sum_{j=1}^{2} x_{ij}(0) = 1 \). To verify this, we first substitute \( \pi_{ij}(x, \mu) \) and \( \overline{\pi}(x, \mu) \) into (3) which yields \( f_{ij}(x, \mu) = \delta \alpha \left( \mu_{ij}(t) k_{ij}(1-p_{ij}) - \beta x_{ij}(t) \sum_{i=1}^{M} C_{i} \right)/N \). Accordingly, we can obtain \( \sum_{i=1}^{M} \sum_{j=1}^{2} f_{ij}(x, \mu) = 0 \) which indicates that the summation of all population shares does not vary with time. Also, we can obtain \( f_{ij}(x, \mu) \geq 0 \) with \( x_{ij}(t) = 0 \) and \( f_{ij}(x, \mu) \leq 0 \) with \( x_{ij}(t) = 1 \). Therefore, \( x_{ij}(t) \in [0, 1] \) can be guaranteed.

\textbf{B. Existence and Uniqueness of Population State Evolution Trajectory under Bandwidth Allocation Control}

We observe that the vector of bandwidth allocation strategies \( \mu \) controls the evolution of the population state. The existence and uniqueness of the solution trajectory to the service selection dynamical system under bandwidth allocation control are discussed in the following theorem.

\textbf{Theorem 1:} For the dynamical system \( \dot{x}_{ij}(t) = f_{ij}(x, \mu), i \in \{1, 2, \ldots, M\}, j \in \{1, 2\} \), with initial condition \( x(0) = x_0 \), if every element in \( \mu \) is a measurable function on \( [0, \infty) \), then there exists a unique solution trajectory \( x(t) \) defined for all \( t \in [0, \infty) \).

\textbf{Proof:} For fixed \( t \), the partial derivative of \( f_{ij}(x, \mu) \) with respect to \( x_{ij}(t) \) is continuous, and if \( \mu_{ij}(t) \) is measurable on \( [0, \infty) \), then \( f_{ij}(x, \mu) \) is measurable for fixed \( x_{ij}(t) \) on the same interval. Moreover, for

\( \text{\footnotesize{Footnote: If the simplicity of presentation, the time parameter is omitted whenever no ambiguity occurs.}} \)

\( \text{\footnotesize{Footnote: For the simplicity of presentation, the time parameter is omitted whenever no ambiguity occurs.}} \)
any given closed bounded set \( K \subseteq \mathbb{R} \) and interval \([a, b] \subseteq [0, \infty)\), there always exists a positive number \( m \) (\( m \geq |x_{ij}(t)| \) for all \( x \in K \)) to construct an integrable function \( g_{ij}(t) = |\delta \alpha_{ij}(t)k_{ij}(1 - p_{ij})/N| + m|\delta \alpha \beta \sum_{l=1}^{M} C_l/N| \) such that \( |f_{ij}(x, \mu)| \leq g_{ij}(t) \) and \( |\partial f_{ij}(x, \mu)/\partial x_{ij}(t)| \leq g_{ij}(t) \) for all \((x, t) \in K \times [a, b]\). Finally, \( |f_{ij}(x, \mu) - f_{ij}(y, \mu)| = \delta \beta \sum_{l=1}^{M} C_l|x_{ij}(t) - y_{ij}(t)|/N \). Denote \( L = \delta \beta \sum_{l=1}^{M} C_l/N + 1 \), we can obtain \( |f_{ij}(x, \mu) - f_{ij}(y, \mu)| \leq L|x_{ij}(t) - y_{ij}(t)| \) which implies that \( f_{ij}(x, \mu) \) satisfies the global Lipschitz condition. Based on the above arguments, we state that the solution to this dynamical system under control is unique and exists globally [30].

\[\text{C. Equilibrium and Evolutionary Stability Analysis}\]

The replicator dynamics describe the variation rate of all population shares. The proportion of AMSs choosing a certain service type provided by an ABS increases if the corresponding payoff is above the average. Evolutionary equilibrium which is defined as the fixed point of replicator dynamics (i.e., \( \dot{x}_{ij}(t) = 0, \forall i, j \)) is considered to be the solution of this service selection evolutionary game. The evolutionary equilibrium can be obtained by solving algebraic equations \( \dot{x}_{ij}(t) = 0, \forall i, j \).

At the equilibrium point, all AMSs in area \( a \) receive the same payoff and none of them has an incentive to change their strategy. Both boundary equilibrium (i.e., \( \exists i, j, x_{ij} = 1 \)) and interior equilibrium (i.e., \( x_{ij} \in (0, 1), \forall i, j \)) exist. For the service selection evolutionary game, the boundary equilibrium is not stable in the sense that any small perturbation will make the system deviate from the equilibrium state. In contrast, we will show below that the interior equilibrium is stable.

We first show that the service selection evolutionary game is a stable population game since the following condition is satisfied for all \( x \neq y \):

\[
\sum_{i=1}^{M} \sum_{j=1}^{2} (x_{ij}(t) - y_{ij}(t))(\pi_{ij}(x, \mu) - \pi_{ij}(y, \mu)) \leq 0. \tag{4}
\]

Therefore, this evolutionary game possesses the property that the equilibrium is neutrally stable [31].

Further, we show that the equilibrium is asymptotically stable. We rewrite (3) in a vector form as \( \dot{x} = Ax + b \), where \( b = [\delta \alpha_{11}k_{11}(1 - p_{11})/N, \delta \alpha_{12}k_{12}(1 - p_{12})/N, \ldots, \delta \alpha C_M \mu_{M2}k_{M2}(1 - p_{M2})/N]^T \), and \( A \) is a \( 2M \times 2M \) dimensional matrix, the characteristic function of which can be expressed as follows:

\[
\det(\gamma I - A) = \left( \gamma + \frac{\delta \alpha \beta}{N} \sum_{l=1}^{M} C_l \right)^{2M} = 0, \tag{5}\]

where \( I \) is an identity matrix. It is straightforward that \( A \) always has \( 2M \) repeated negative eigenvalues which verifies the asymptotically stable property of the service selection dynamics. Therefore,
this dynamical system asymptotically converges to the evolutionary equilibrium point from any initial interior population state \( x_0 \). According to [16], a strategy profile is evolutionary stable if and only if it is asymptotically stable. Therefore, the interior equilibrium is an evolutionary stable strategy (ESS) profile.

V. DYNAMIC BANDWIDTH ALLOCATION IN IEEE 802.16M MOBILE RELAY NETWORKS

With the dynamic service selection behavior of AMSs and queue dynamics, the ABSs competitively allocate the downlink bandwidth for direct transmission and relay-cooperation transmission to maintain the desired queue level and user distribution for satisfying their own system performance requirements. Allocating small bandwidth for a relay-cooperation transmission will waste part of the bandwidth of ARSs while allocating large bandwidth for a relay-cooperation transmission may result in performance degradation of direct transmission. Therefore, the bandwidth allocation needs to be performed optimally under competition. In this section, we formulate the dynamic bandwidth allocation in IEEE 802.16m mobile relay networks as a non-cooperative differential game. This is the upper-level game in the proposed two-level game framework which takes the dynamic service selection of AMSs and the queue state dynamics at the ABSs into account.

A. Dynamic Bandwidth Allocation as a Differential Game

We first formulate the competitive dynamic bandwidth allocation in IEEE 802.16m mobile relay networks as a differential game. In this case, each of the \( M \) ABSs competes to minimize its own cost function (to be explained) by controlling the bandwidth allocation strategy. The players are the \( M \) ABSs. The strategy of ABS \( i \) is the dynamic control of the downlink bandwidth allocated for the direct transmission and relay-cooperation transmission which can be denoted by \( \mu_i(t) = [\mu_{i1}(t), \mu_{i2}(t)]^T \). Similar to the common notation used in game theory, \( \Phi = \{\mu_i(t), \mu_{-i}(t)\} \) denotes the strategy profile of this differential game, where \( \mu_{-i}(t) \) is a vector of strategies of all ABSs except ABS \( i \).

Depending on different informational structures, the control strategies can be represented in the different ways (i.e., open-loop and closed-loop). An open-loop strategy \( \mu_i = \mu_{i\text{ol}}(t) \) is a function of time alone which does not need any feedback information (e.g., system state). With open-loop strategies, each player determines the action plan in advance and complies with the plan during the play. Also, the output of the control does not need to be measured. In comparison, a closed-loop strategy \( \mu_i = \mu_{i\text{cl}}(x, t) \) can use feedback information to adjust the control process. The player adopting a closed-loop strategy commits to
the control law instead of a particular control path. Closed-loop control has many advantages over open-loop control (e.g., stabilizing unstable systems and handling disturbance). Closed-loop control needs to estimate the system state from the observed output which adds complexity to the system design. In general, for wireless networks, the closed-loop control can be implemented in infrastructure-based networks with the central controller collecting and distributing the state information, while the open-loop control is suitable for ad hoc networks for which the state measurement and information distribution are costly. In this paper, we consider the closed-loop control due to its advantages over open-loop control strategy for centralized network architecture of IEEE 802.16m. As common in control theory, the time parameter will be omitted whenever no ambiguity occurs.

In the bandwidth allocation differential game, the control strategy of each ABS influences the state evolution of the system as well as the values of cost function of other ABSs. The population state $x$ of the underlying service selection evolutionary game of AMSs and the queue state $q$ of the ABSs jointly describe the system state of the upper-level differential game. With $n_{ij} = N x_{ij}$ and $r = r_n + \xi$, the queue dynamics defined in (1) can be rewritten as follows:

$$
\dot{q}_{ij} = N x_{ij} r_n - (1 - p_{ij}) \mu_{ij} k_{ij} + N x_{ij} \xi, \quad (6)
$$

for all $i \in \{1, 2, \ldots, M\}$ and $j \in \{1, 2\}$. The replicator dynamics (3) and queue dynamics (6) describe how the current population state $x$, queue state $q$, bandwidth allocation strategies $\Phi$, and the disturbance $\xi$ influence the rate of change of the system state. Note that the multiplicative disturbance appearing in the queue dynamics indicates that the more subscribed users, the larger uncertainties introduced to the system by the disturbance.

The objective of the dynamic bandwidth allocation can be described from a system performance point of view. First, each ABS tries to maintain the queue length as close as possible to the target level for satisfying delay requirements with minimum required bandwidth. Second, each ABS tries to drive the AMS population state to the target distribution to avoid wasting bandwidth as well as overloading system. Third, each ABS tries to prevent large disparities between aggregated traffic input flow rate and the allocated bandwidth. For measuring the performance, a quadratic instantaneous cost function is considered as follows:

$$
J_{\text{ins}}(\mu_i, \mu_{-i}) = \sum_{j=1}^{2} \left[ w_{ij1} (q_{ij} - Q_{ij}^{\text{tar}})^2 + w_{ij2} (N x_{ij} - N_{ij}^{\text{tar}})^2 + w_{ij3} (N x_{ij} r - \mu_{ij} k_{ij} (1 - p_{ij}))^2 \right], \quad (7)
$$
where $Q_{ij}^{tar}$ and $N_{ij}^{tar}$ are the target queue length and target number of AMSs, respectively. The first two terms (i.e., $w_{ij1}(q_{ij} - Q_{ij}^{tar})^2 + w_{ij2}(N_{ij} - N_{ij}^{tar})^2$) in the cost function describe the penalties for deviating from the target queue length and target population distribution, respectively. The third term (i.e., $w_{ij3}(N_{ij} - \mu_{ij}(1 - p_{ij}))^2$) measures the disparities among the aggregated traffic input flow rate and the allocated system bandwidth. $w_{ij1}$, $w_{ij2}$, and $w_{ij3}$ are positive weighting coefficients reflecting the relative importance of the requirements as well as the service priority.

For the formulated non-cooperative bandwidth allocation differential game, the closed-loop Nash equilibrium is a reasonable notion of optimality and is considered to be the solution. The definition of the closed-loop Nash equilibrium is given as follows:

**Definition 1:** Denote $J_i(\mu_i, \mu_{-i})$ the cost function of ABS $i$. The closed-loop strategy profile $\{\mu_i^*, \mu_{-i}^*\}$ is a closed-loop Nash equilibrium if for each ABS $i \in \{1, 2, \ldots, M\}$, $J_i(\mu_i^*, \mu_{-i}^*) \leq J_i(\mu_i, \mu_{-i}^*)$ holds for all feasible control paths $\mu_i$ in the bandwidth allocation differential game given other ABSs’ control strategies $\mu_{-i}^*$.

To obtain the closed-loop Nash equilibrium of this differential game, each ABS needs to solve an optimal control problem under disturbance. In this case, two approaches (i.e., stochastic optimal control and $H_{\infty}$ robust optimal control) can be used depending on the characteristics of the disturbance $\xi$. If $\xi$ enters the system dynamics in the form of white noise Gaussian process, the stochastic optimal control approach can be used to design a controller optimizing the average performance. If $\xi$ is energy bounded, while the statistics are unknown, the $H_{\infty}$ robust optimal control approach can be used for the worst-case performance control design.

**B. Stochastic Optimal Control Approach**

We first consider the stochastic optimal control approach with the assumption that the disturbance $\xi$ is a white noise Gaussian process. In this case, the system state evolution is subject to a white noise disturbance and described by stochastic differential equations. For each of $M$ ABSs, the system performance optimization becomes a stochastic optimal control problem subject to the constraints (i.e., stochastic state evolution differential equations and physical constraints) given the control strategies of other ABSs. The stochastic optimal control problem can be expressed as follows:

$$\min_{\mu_i} J_i(\mu_i, \mu_{-i}) = \min_{\mu_i} \lim_{T \to \infty} \frac{1}{T} E \left\{ \int_0^T J_{ins}(\mu_i, \mu_{-i}) dt \right\},$$  

(8)
respectively. In this case, the system dynamics can be written in a vector form as follows:

\[
\begin{align*}
\dot{x}_{ij} &= \delta x_{ij} \left( \pi_{ij}(x, \mu) - \bar{\pi}(x, \mu) \right), \\
\dot{q}_{ij} &= N x_{ij} r_n - (1 - p_{ij}) \mu_{ij} k_{ij} + N x_{ij} \xi,
\end{align*}
\]

(9)

for all \(i \in \{1, 2, \ldots, M\}, j \in \{1, 2\}\). Note that the state evolution cannot be fully determined given the current state and control path since the dynamics are stochastic. Therefore, the minimization is taken over the expectation value over all possible future realizations of the white noise process.

We first rewrite the population dynamics and queue dynamics (9) as follows:

\[
\begin{align*}
\hat{x}_{ij} &= \delta \alpha \left( - \sum_i M C_i / N + r_n \right) \bar{x}_{ij} - \delta \alpha / N \bar{u}_{ij} + \delta \alpha N_{ij}^\text{tar} (r_n / N - 1) + \delta \alpha (\bar{x}_{ij} + N_{ij}^\text{tar} / N) \xi, \\
\hat{q}_{ij} &= N x_{ij} r_n - (1 - p_{ij}) \mu_{ij} k_{ij}.
\end{align*}
\]

(10)

where \(\hat{x}_{ij} = x_{ij} - N_{ij}^\text{tar} / N\), \(\hat{q}_{ij} = q_{ij} - Q_{ij}^\text{tar}\), and \(\bar{u}_{ij} = N x_{ij} r_n - (1 - p_{ij}) \mu_{ij} k_{ij}\). We now define \(\tilde{y}_i = [\tilde{x}_{i1}, \tilde{x}_{i2}, \tilde{q}_{i1}, \tilde{q}_{i2}]^T\) and \(\tilde{u}_i = [\tilde{u}_{i1}, \tilde{u}_{i2}]^T\) which serve as the new vectors of system state and control strategies, respectively. In this case, the system dynamics can be written in a vector form as follows:

\[
d\tilde{y}_i = [\tilde{A}_i \tilde{y}_i + \tilde{B}_i \tilde{u}_i + \tilde{d}_i]dt + [\tilde{E}_i \tilde{y}_i + \tilde{d}_i]d\xi,
\]

(11)

where the coefficient matrices \(\tilde{A}_i, \tilde{B}_i, \tilde{E}_i, \tilde{d}_i\), and \(\tilde{d}_i\) with suitable dimensions can be derived according to (10). The instantaneous cost function for provider \(i\) can be written as follows:

\[
J_{\text{ins}}^i(\tilde{u}_i, \tilde{y}_i) = \tilde{y}_i^T \tilde{Q}_i \tilde{y}_i + \tilde{u}_i^T \tilde{R}_i \tilde{u}_i,
\]

(12)

where \(\tilde{Q}_i\) and \(\tilde{R}_i\) are positive definite matrices which can be derived from (7).

To obtain a closed-loop Nash equilibrium, Hamilton Jacobi Bellman (HJB) method [32] can be used. As common for autonomous dynamic games with infinite time horizon, the aim is to find stationary Nash equilibrium. In this case, the optimal cost function and optimal control strategy are implicit of time \(t\). Specifically, denote \(V_i(\tilde{y}_i)\) the optimal cost function for ABS \(i\). The HJB equation for the bandwidth allocation differential game under white noise stochastic disturbance is expressed as follows:

\[
-\frac{1}{2} \text{Tr} \left( \frac{\partial^2 V_i(\tilde{y}_i)}{\partial \tilde{y}_i^2} \sigma(\tilde{y}_i)^T \sigma(\tilde{y}_i) \right) = \min_{\tilde{u}_i} \left\{ J_{\text{ins}}^i(\tilde{u}_i, \tilde{y}_i) + \frac{\partial V_i(\tilde{y}_i)}{\partial \tilde{y}_i} g_i(\tilde{u}_i, \tilde{y}_i) \right\},
\]

(13)

where \(\text{Tr}(\cdot)\) is the trace of a matrix, \(\sigma(\tilde{y}_i) = \tilde{E}_i \tilde{y}_i + \tilde{d}_i\), and \(g_i(\tilde{u}_i, \tilde{y}_i) = \tilde{A}_i \tilde{y}_i + \tilde{B}_i \tilde{u}_i + \tilde{d}_i\).

The optimal control strategy \(\tilde{u}_i^*\) of ABS \(i\) minimizes the right hand side of (13) and can be expressed as a function of \(V_i(\tilde{y}_i)\). In our case, the state dynamics are linear and the cost is quadratic. Therefore the optimal cost function is also a quadratic form denoted by

\[
V_i(\tilde{y}_i) = \frac{1}{2} \tilde{y}_i^T \tilde{P}_i \tilde{y}_i + \tilde{p}_i \tilde{y}_i + \tilde{b}_i,
\]

(14)
where \( \tilde{P}_i, \tilde{\phi}_i, \) and \( \tilde{\beta}_i \) are coefficient matrices needed to be determined. Substituting (14) into (13) and doing the minimization with respect to \( \tilde{u}_i \) yields

\[
\tilde{u}_i^* = -\tilde{\Psi}_i \tilde{y}_i - \tilde{\psi}_i,
\]

where \( \tilde{\Psi}_i = \tilde{R}_i^{-1} \tilde{B}_i^T \tilde{P}_i \) and \( \tilde{\psi}_i = \tilde{R}_i^{-1} \tilde{B}_i^T \tilde{\varphi}_i \). Substituting (14) and (15) into (13) and collecting terms with the same power yields the following algebraic equations:

\[
\tilde{P}_i \tilde{A}_i + \tilde{A}_i^T \tilde{P}_i + \tilde{E}_i^T \tilde{P}_i \tilde{E}_i + \tilde{Q}_i - (\tilde{B}_i^T \tilde{P}_i)^T \tilde{R}_i^{-1} (\tilde{B}_i^T \tilde{P}_i) = 0,
\]

\[
\frac{1}{2} \left( \sqrt{\tilde{R}_i \tilde{\psi}_i} \right)^2 - \tilde{\varphi}_i \tilde{b}_i - \frac{1}{2} \tilde{d}_i^T \tilde{P}_i \tilde{d}_i = 0.
\]

\( \tilde{P}_i \) is determined by solving the algebraic Riccati equation for which the linear matrix inequality (LMI) can be used. The problem is equivalent to \( \tilde{P}_i = \arg \min_{\tilde{P}_i} \text{Tr}(\tilde{P}_i) \) subject to

\[
\begin{bmatrix}
\tilde{P}_i \tilde{A}_i + \tilde{A}_i^T \tilde{P}_i + \tilde{E}_i^T \tilde{P}_i \tilde{E}_i + \tilde{Q}_i - (\tilde{B}_i^T \tilde{P}_i)^T \tilde{R}_i^{-1} (\tilde{B}_i^T \tilde{P}_i) & \tilde{B}_i^T \tilde{P}_i \\
\tilde{B}_i \tilde{P}_i & -I
\end{bmatrix} \leq 0.
\]

In this case, \( \tilde{P}_i \) can be determined, and then \( \tilde{\varphi}_i, \tilde{\Psi}_i, \) and \( \tilde{\psi}_i \) can be obtained accordingly.

We have obtained the optimal value function as well as the closed-form optimal bandwidth allocation strategy \( \tilde{u}_i^* \) for ABS \( i \). Similarly, the optimal control strategies for all other ABSs can be obtained. Accordingly, the strategy profile \( \Phi^* = \{ \tilde{u}_i^*, \tilde{u}_i^{*-1} \} \) constitutes the closed-loop Nash equilibrium for the bandwidth allocation differential game under white noise traffic input flow rate stochastic disturbance.

**C. \( H_\infty \) Robust Optimal Control Approach**

With stochastic optimal control approach, we have the assumption that the traffic input flow rate disturbance, though unpredictable, is not completely unknown. The statistics of the disturbance are known as a priori knowledge. For the most general cases where the disturbance statistics cannot be adequately described, \( H_\infty \) robust optimal control as an alternative approach can be used to provide the worst-case performance guarantee which makes no assumption on the disturbance \( \xi \) other than it is constrained to have finite energy in terms of bounded \( L_2 \) norm (e.g., continuous deterministic damped traffic input flow rate disturbance). As one of the robust control techniques, \( H_\infty \) optimal control explicitly takes the system model (or the noise model) uncertainties into account. The objective is to obtain a controller design that minimizes the performance cost function under the worst possible disturbances.
For the derivation of $H_\infty$ robust optimal controller for the dynamic bandwidth allocation which minimizes the $H_\infty$ norm of the system, the system state and cost function need to be redefined. First, we introduce a new state variable $\hat{y}_i = [\tilde{x}_{i1}, \tilde{x}_{i2}, \tilde{q}_{i1}, \tilde{q}_{i2}, e]$ where $e = 1$ is a constant. In this case, the system state dynamics and instantaneous objective function can be respectively rewritten as follows:

\[
\dot{\hat{y}}_i = \hat{A}_i\hat{y}_i + \hat{B}_i\hat{u}_i + \hat{E}_i\xi, \quad \text{and} \quad J_{\text{ins}}^{i}(\hat{u}_i, \hat{y}_i) = \hat{y}_i^T\hat{Q}_i\hat{y}_i + \hat{u}_i^T\hat{R}_i\hat{u}_i,
\]

where $\hat{u}_i = \tilde{u}_i$, $\hat{A}_i$, $\hat{B}_i$, $\hat{C}_i$, $\hat{Q}_i$, and $\hat{R}_i$ can be constructed according to (10). Since the state variable (e.g., queue length) can be internally obtained from the ABS, this system is classified as continuous-time with perfect state measurements.

In the following, the cost function as dynamic bandwidth allocation performance index in the context of $H_\infty$ optimal control will be introduced. First, the controlled output $z_i$ of ABS $i$ is defined as follows:

\[
z_i = \hat{H}_i\hat{y} + \hat{G}_i\hat{u}_i,
\]

where matrices $\hat{H}_i$ and $\hat{G}_i$ represent the weights of cost for deviations from target state (e.g., queue length) and actual aggregated downlink traffic flow rate, respectively. In particular, we have

\[
\begin{bmatrix}
\hat{H}_i^T\hat{H}_i & \hat{H}_i^T\hat{G}_i \\
\hat{G}_i^T\hat{H}_i & \hat{G}_i^T\hat{G}_i
\end{bmatrix} = \begin{bmatrix}
\hat{Q}_i & 0 \\
0 & \hat{R}_i
\end{bmatrix},
\]

(18)

where

\[
\hat{Q}_i = \begin{bmatrix}
\hat{Q}_i & 0 \\
0 & 0
\end{bmatrix}, \quad \hat{R}_i = \begin{bmatrix}
\hat{R}_i & 0 \\
0 & 0
\end{bmatrix}.
\]

(19)

The term $\hat{H}_i^T\hat{G}_i = 0$ indicates the absence of cross terms between the state $\hat{y}_i$ and control $\hat{u}_i$ in the cost function of ABS $i$. Next, the $L_2$ norm of the controlled output $z_i$ is defined as follows:

\[
\|z_i\|_2 = \left(\int_0^\infty z_i^Tz_i\,dt\right)^{\frac{1}{2}}.
\]

(20)

Accordingly, the $H_\infty$ norm of system which reflects the worst-case $L_2$ gain of the system due to the exogenous signal $\xi$ is defined as follows:

\[
\|L_i\|_\infty = \sup_{\xi} \frac{\|z_i\|_2}{\|\xi\|_2}.
\]

(21)

With the definition of the $H_\infty$ norm, the robust optimal controller design problem for the dynamic bandwidth allocation of ABSs is expressed as a mathematical optimization problem which is concerned with finding a feasible state feedback controller $\hat{u}_i = \hat{K}_i\hat{y}_i$ such that the $H_\infty$ norm of the system is minimized, i.e.,

\[
\min_{\hat{u}_i} \sup_{\xi} \frac{\|z_i\|_2}{\|\xi\|_2}.
\]

(22)
From (22), we can see that the $H_\infty$ optimal control problem is in fact a minmax optimization problem. A natural way to solve this problem is to formulate a zero-sum sub-differential game in which the controller can be viewed as the minimizing player and the disturbance as the maximizing player [33]. The $H_\infty$ norm cannot be used as the cost function directly for the sub-differential game since the solution of which is intractable in general. Instead, as common in $H_\infty$ optimal control theory, an augmented parameterized cost function is used and defined as follows: $J = \|z\|^2 - \gamma^2 \|\xi\|^2$, where $\gamma > 0$ is a bound on the $H_\infty$ norm defined as follows:

$$\|L_i\|_\infty = \sup_\xi \frac{\|z\|^2}{\|\xi\|^2} < \gamma.$$  

(23)

Note that this augmented cost function is introduced for purely technical reasons. The scalar $\gamma$ is a prescribed value interpreted as the performance bound which measures the ability of the system to mitigate the effect of the disturbance $\xi$ on the output. A controller that satisfies this bound is called the suboptimal solution to the $H_\infty$ optimal control problem. The performance of the controller relates to the value of $\gamma$. In particular, larger $\gamma$ results in more conservative controller.

For the sub-differential game, the controller seeks to minimize the cost $J$ while the disturbance viewed as another player seeks to maximize the cost $J$. The solution of this differential game is necessarily a saddle point of the cost function, which is defined by the following inequalities

$$J(\hat{u}_i^*, \xi) \leq J(\hat{u}_i^*, \xi^*) \leq J(\hat{u}_i, \xi^*).$$

(24)

The suboptimal control can be obtained as $\hat{u}_i^* = -\hat{B}_i \hat{P}_i \hat{y}_i$, where $\hat{P}_i$ is the minimal semi-definite solution of the following algebraic Riccati equation

$$\hat{P}_i \hat{A}_i + \hat{A}_i^T \hat{P}_i + \hat{H}_i^T \hat{H}_i - \hat{P}_i (\hat{B}_i \hat{B}_i^T - \gamma^2 \hat{E}_i \hat{E}_i^T) \hat{P}_i = 0.$$  

(25)

Linear matrix inequality (LMI) can also be used for solving this Riccati equation. The $H_\infty$ optimal controller can be obtained with the minimal $\gamma$ and it can be approximated numerically by decreasing the performance bound until the feasible solution of (25) no longer exists.

With the obtained robust optimal bandwidth allocation strategies $\hat{u}_i^*$ for all ABSs, the strategy profile $\Phi^* = \{\hat{u}_i^*, \hat{u}_{-i}^*\}$ constitutes the closed-loop Nash equilibrium for the bandwidth allocation differential game under general energy bounded input rate disturbance. The detailed implementation schemes for the stochastic optimal controller and $H_\infty$ optimal controller are provided in Table I.
Note that the detailed design of a specific protocol for the game framework is out of the scope of this paper. However, the information required for the implementation of algorithms can be analyzed. Specifically, each ABS needs to compute the average population payoff according to the user distribution and then distributes it to the subscribed AMSs. Besides, for the measurement of population distribution, each ABS needs to exchange the population distribution information with other ABSs.

VI. PERFORMANCE EVALUATION

A. Parameters Setting

We consider a competitive IEEE 802.16m mobile relay network with two ABSs (i.e., \( M = 2 \)) providing access service to 150 active AMSs (i.e., \( N = 150 \)) in a particular service area. The total downlink bandwidth of the ABSs 1 and 2 for this area are \( C_1 = 15 \text{ MHz} \), and \( C_2 = 10 \text{ MHz} \), respectively. The average spectrum efficiency is assumed to be 1.2 bits/s/Hz. The packet error rate are assumed to be \( p_{11} = p_{21} = 0.02 \) and \( p_{12} = p_{22} = 0.04 \). The nominal traffic input flow rate is assumed to be \( r_n = 180 \text{ kbps} \). For the replicator dynamics, we set the learning rate to be \( \delta = 1 \). For the utility function, we set \( \alpha = \beta = 1 \). We consider equal importance of different target requirements as well as service priority and we set \( w_{ij1} = w_{ij2} = w_{ij3} = 1 \) \( \forall i, j \). The initial proportion of AMSs choosing two service types of two ABSs is assumed to be \( x_{11}(0) = 0.2, \ x_{12}(0) = 0.3, \ x_{21}(0) = 0.1, \) and \( x_{22}(0) = 0.4 \). The initial backlogged data are assumed to be \( q_{11}(0) = 150\text{kb}, \ q_{12}(0) = 130\text{kb}, \ q_{21}(0) = 120\text{kb}, \) and \( q_{22}(0) = 100\text{kb} \). The target queue lengths are set to be \( Q_{11}^{\text{tar}} = Q_{12}^{\text{tar}} = Q_{21}^{\text{tar}} = Q_{22}^{\text{tar}} = 40\text{kb} \). The target user distributions are set to be \( N_{11}^{\text{tar}} = 60, \ N_{12}^{\text{tar}} = 30, \ N_{21}^{\text{tar}} = 40, \) and \( N_{22}^{\text{tar}} = 20 \).

B. Numerical Results

The dynamic service selection behavior of AMSs subject to the bandwidth allocation control of ABSs under Gaussian white noise is investigated and the sample population share trajectories from the initial selection distribution are shown in Fig. 3. When the proportion of AMSs selecting a certain transmission mode of an ABS is higher than the equilibrium proportion, the AMSs in this group receive less utility than the average. As a result, some AMSs will switch to other strategies to achieve better performance. Fig. 3 shows that the state evolution paths converge to a certain population distribution at which all AMSs receive the same average utility of the population. To accommodate the traffic input flow rate fluctuation, the bandwidth allocated by closed-loop controls is dynamic and fluctuated (Fig. 4) which results in a
fluctuated service selection distribution. In this case, the actual state evolution path fluctuates around the mean path but never remains at the steady state. Fig. 5 shows the empirical mean and 95% confidence interval of the population share evolution path \( x_{11} \). The interval provides a precise bound for the fluctuated population share due to disturbance.

We obtain the closed-loop Nash equilibrium control strategies of the bandwidth allocation differential game based on stochastic optimal control and \( H_\infty \) optimal control for optimizing average performance and worst-case performance, respectively. In addition, for comparison purpose, we formulate a static non-cooperative bandwidth allocation game for which the static Nash equilibrium control strategy is obtained only considering the equilibrium distribution of the underlying evolutionary game. The performance of different control strategies for maintaining the queue length is evaluated in two cases. In Case I, the traffic input flow rate is random and follows Gaussian distribution \( \mathcal{N}(0, \sigma^2) \). In Case II, energy bounded deterministic continuous disturbance and jitters (impulse traffic) occur on the traffic input flow rate. Specifically, we assume

\[
R = \begin{cases} 
  r_n + 50e^{-0.01t} + 50, & 200 < t < 400, 1000 < t < 1200 \\
  r_n + 50e^{-0.01t}, & \text{otherwise}.
\end{cases}
\] (26)

First, the queue dynamics under different control strategies of Case I is investigated. The queue evolution of \( q_{11} \) is shown in Fig. 6 as an example. All three strategies control the queues from initial states to a level around target length. The stochastic optimal control approach performs the best in terms of convergence speed. In contrast, the static control gives the worst performance. Also, due to continuously stochastic disturbance, the queue dynamics fluctuate around certain queue level.

Next, we vary the system parameters and investigate the impact on system performance. As shown in Fig. 7, the larger disturbance variance, the larger fluctuation to the queue dynamics is introduced. Also, with the increasing mean of the disturbance, the aggregated data rate increases and the network becomes overloaded. Therefore, the queue dynamics converge to a level above the target level.

The queue dynamics of \( q_{11} \) under different control strategies of Case II are investigated in Fig. 8. In this case, the disturbance on the traffic input flow rate is norm-bounded and the Gaussian white noise assumption is no longer valid. As shown in Fig. 8, all three control strategies respond to the continuous disturbance and the traffic flow rate jitters. The queue length increases when the jitter occurs. In Case II, \( H_\infty \) optimal control approach performs the best following by the stochastic optimal control approach, while the static control still gives the worst performance.
In summary, if the traffic input flow rate disturbance is white noise (Case I), the stochastic optimal controller has the better performance since it is optimal in the sense that it minimizes the expected value of the performance. However, if under general energy bounded disturbance (Case II), the $H_\infty$ optimal controller performs better due to its robustness of the worst-case design. Both these two dynamic controls have evident performance enhancement compared with static control in both cases.

VII. CONCLUSION

In this paper, we have developed a two-level dynamic game framework to jointly address the problem of dynamic service selection and dynamic bandwidth allocation in IEEE 802.16m mobile relay networks. The evolution and dynamic service selection behavior of advanced mobile stations have been modeled as an evolutionary game and the strategy adaption process has been analyzed using the replicator dynamics. The dynamic competition among providers considering mobile stations’ service selection and queue states of the advanced base stations has been formulated as a differential game. The service selection distribution and the queue states jointly describe the state of the differential game. Both stochastic optimal control and $H_\infty$ robust optimal control are used to obtain the closed-loop Nash equilibrium which is considered to be the solution of the differential game depending on the different disturbance types. Numerical studies have been performed to investigate the effectiveness of dynamic controls and to compare the system performance under different control strategies. For the future work, the cooperation among different providers and a non-autonomous system can be investigated. Also, due to the control signal delay in the uplink direction, a time-delayed uplink bandwidth allocation framework can be studied.

REFERENCES


Fig. 1. The frame structure of IEEE 802.16m.

Fig. 2. System model of a competitive IEEE 802.16m mobile relay network with $M$ advanced base stations and $N$ advanced mobile stations.
Fig. 3. Population state of the service selection evolutionary game under stochastic optimal control.

Fig. 4. Service rate of different queues under stochastic optimal control.

Fig. 5. Empirical mean and 95% confidence interval (CI) under stochastic optimal control.
Fig. 6. Queue states of $q_{11}$ under Gaussian noise.

Fig. 7. Queue states of the $q_{11}$ under stochastic disturbance with different mean and variance.

Fig. 8. Queue states of $q_{11}$ under non-Gaussian noise.
<table>
<thead>
<tr>
<th>Control strategies</th>
<th>Implementation schemes</th>
</tr>
</thead>
</table>
| **Stochastic optimal controller** | **For each ABS $i$ do,**  \[\text{For each ABS } i \text{ do,}\]  
Measure the population distribution $\mathbf{x}$ and queue state $\mathbf{q}_i$;  
Update $\hat{\mathbf{u}}_i$ using (15);  
Update the bandwidth allocation strategy $\mu_i$ at each decision point;  
**End for**                                                                 |
| **$H_\infty$ optimal controller** | **For each ABS $i$ do,**  \[\text{For each ABS } i \text{ do,}\]  
Measure the population distribution $\mathbf{x}$ and queue state $\mathbf{q}_i$;  
Update $\hat{\mathbf{u}}_i$ using $\hat{\mathbf{u}}_i^* = \hat{\mathbf{K}}_i \hat{\mathbf{y}}_i$;  
Update the bandwidth allocation strategy $\mu_i$ at each decision point;  
**End for** |