Liquid State Machine Learning for Resource Allocation in a Network of Cache-Enabled LTE-U UAVs

Mingzhe Chen*, Walid Saad†, and Changchuan Yin*
*Beijing Key Laboratory of Network System Architecture and Convergence,
Beijing University of Posts and Telecommunications, Beijing, China 100876
Emails: chenmingzhe@bupt.edu.cn, and ccyin@bupt.edu.cn.
†Wireless@VT, Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA. Email: walids@vt.edu.

***Abstract*—In this paper, the problem of joint caching and resource allocation is investigated for a network of cache-enabled unmanned aerial vehicles (UAVs) that service wireless ground users over the LTE licensed and unlicensed (LTE-U) bands. The considered model focuses on users that can access both licensed and unlicensed bands while receiving contents through UAV cache-user links and content server-UAV-user links. This problem is formulated as an optimization problem which jointly incorporates user association, spectrum allocation, and content caching. To solve this problem, a distributed algorithm based on the machine learning framework of *liquid state machine* (LSM) is proposed. Using the proposed LSM algorithm, the cloud can predict the users' content request distribution while having only limited information on the network's and users' states. The proposed algorithm also enables the UAVs to autonomously choose the optimal resource allocation strategies depending on the network states. Simulation results using real datasets show that the proposed approach yields up to 33.3% and 50.3% gains, respectively, in terms of the number of users that have stable queues compared to two baseline algorithms: Q-learning with cache and Q-learning without cache. The results also show that LSM significantly improves the convergence time of up to 33.3% compared to Q-learning.**

I. Introduction

The use of aerial wireless communication platforms such as unmanned aerial vehicles (UAVs) is seen as a promising approach to improve the coverage and capacity of future wireless networks [1]. Due to their flying nature, UAVs can provide line-of-sight (LoS) connections toward ground users and, thus, potentially improving the rate, delay, and overall performance of wireless networks. However, deploying UAVs for wireless communication purposes faces many challenges [2]–[5] that include air-to-ground channel modeling, optimal deployment, energy efficiency, path planning, and resource management.

The existing literature in [3]–[5] has studied a number of problems related to UAVs. In [3], the deployment of an unmanned aerial vehicle acting as wireless base stations that provide coverage for ground users is analyzed. The authors in [4] propose a statistical propagation model for predicting the air-to-ground path loss between a low altitude platform and a ground user. In [5], the authors investigate a multi-tier drone-enabled network complementing a terrestrial heterogeneous network. However, most of these existing works are focused on the UAV-ground users wireless communications and do not account for the ground-to-air fronthaul communications between the core network and the UAVs. Indeed, capacity-constrained fronthaul links will significantly limit the transmission rate of the links from the UAVs to the ground users. In order to reduce the traffic load on the fronthaul links and improve the performance of UAV-based communication, one promising approach is to cache popular content [6] at the UAVs thus allowing them to directly transmit data to the ground users without using wireless fronthaul transmissions [7].

To further improve performance and overcome the spectrum scarcity problem, the UAVs can be equipped with LTE over the unlicensed band (LTE-U) capabilities thus allowing them to use both licensed and unlicensed spectrum to service their ground users. Recently, there has been significant interest in studying the performance of LTE-U enabled cellular networks such as in [8]–[10]. In [8], the authors investigate the use of LTE-U for unmanned aerial base stations to enhance the achievable broadband throughput during emergency situations. The authors in [9] introduce a deep learning approach for resource allocation problem in LTE-U SCNs. In [10], the authors propose a novel learning algorithm to solve the problem of resource allocation with uplink-downlink decoupling. However, the LTE-U works in [9] and [10] do not consider the use of LTE-U with UAV-carried base stations. Meanwhile, the work in [8] considers the LTE-U resource allocation in a UAV network, however, it does not consider the effect of the limited-capacity cloud-UAVs links. Caching the popular contents at UAVs can help overcome the limited capacity of the cloud-UAVs links.

The main contribution of this paper is a novel resource allocation framework for allowing cache-enabled UAVs to effectively service ground users over licensed and unlicensed bands in a cloud network under fronthaul capacity constraints. The proposed approach will enable *dual-mode* UAVs to autonomously learn and determine which content to cache and how to allocate the licensed and unlicensed bands to each user depending on the network environment. Developing such a dynamic resource allocation algorithm requires a self-organizing, decentralized approach so as to minimize the overhead and coordination among UAVs while maximizing performance. Unlike previous studies [3]–[5], that overlook the limited-capacity of the UAV-cloud links, we propose a novel learning approach based on the powerful framework of *liquid state machine (LSM)* [11] to perform caching and resource allocation in a network of cache-enabled LTE-U UAVs. The use of LSM enables the cloud to quickly learn the users' content request distribution so as to determine the content caching strategy for each UAV. It also enables the UAVs to autonomously adjust their spectrum allocation schemes to service users. To our best knowledge, *this is the first work to jointly consider the use of the caching and LTE-U for UAV-assisted communication*. Simulation results show that the proposed approach yields, respectively, 33.3% and 50.3% gains in terms of the number of users having a stable queue compared to Q-learning with cache and Q-learning without cache.

The rest of this paper is organized as follows. The system

This work was supported in part by the National Natural Science Foundation of China under Grants 61671086 and 61629101, in part by BUPT-SICE Excellent Graduate Students Innovation Fund, and in part by the Office of Naval Research (ONR) under Grant CNS-1513697, and, in part, by the National Science Foundation under Grants IIS-1633363 and AST-1506297.

model and problem formulation are described in Section II. The LSM-based algorithms for content request distribution prediction and resource allocation are proposed in Section III. In Section IV, numerical simulation results are presented and analyzed. Finally, conclusions are drawn in Section V.

II. System Model and problem formulation

Consider the downlink of an LTE-U network composed of a set $\mathcal{K}$ of K UAVs and W WiFi access points (WAPs). In this model, the UAVs are equipped with cache storage units [12] and can be deployed to act as flying cache-enabled LTE-U base stations to serve a set $\mathcal{U}$ of U ground users. The UAVs are controlled by a cloud-based server. Here, we consider *dual mode* cache-enabled UAVs that are able to access both the licensed and unlicensed bands. The transmissions from the cloud to the UAVs occur over wireless fronthaul links using the licensed cellular band.

In this system, a frequency division duplexing (FDD) mode is considered for LTE on the licensed band. The FDD mode separates the licensed band for the downlink LTE-U users. A time division duplexing (TDD) mode with duty cycle method is considered for LTE-U. Using the duty cycle method, the UAVs will use a discontinuous, duty-cycle transmission pattern so as to guarantee the transmission rate of WiFi users. Under this method, the unlicensed band time slots will be divided between LTE-U and WiFi users. In particular, LTE-U transmits for a fraction ϑ of time and will be muted for $1-\vartheta$ time which is allocated for WiFi transmission. The WAPs transmit using a standard carrier sense multiple access with collision avoidance (CSMA/CA) protocol and its corresponding RTS/CTS access mechanism.

In our model, we assume that all of the users will only request contents of equal size L from a set $\mathcal{N}$ of N contents that are stored at a cloud-based content server. Each UAV k is equipped with a storage unit that can store a set $\mathcal{C}_k$ of C popular contents that the users can request. Caching at the UAVs can significantly offload the fronthaul traffic of UAVs since each UAV can directly transmit its stored contents to the users without using fronthaul links. Hereinafter, caching at the UAVs is referred to as "UAV cache". The cached contents at a UAV are assumed to be refreshed at off peak hours when the UAVs return to their docking stations.

A. WiFi data rate analysis

For the WiFi network, we assume that the WAPs will adopt a CSMA/CA scheme with binary slotted exponential backoff. Therefore, the saturation capacity of N_w users sharing the same unlicensed band can be expressed by [10]:

$$R\left(N_w\right) = \frac{P_{\text{tr}}\left(N_w\right) P_{\text{s}}\left(N_w\right) E\left[A\right]}{(1-P_{\text{tr}}\left(N_w\right))T_\sigma + P_{\text{tr}}(N_w) P_{\text{s}}(N_w) T_{\text{s}} + P_{\text{tr}}(N_w)(1-P_{\text{s}}(N_w)) T_{\text{c}}}, \tag{1}$$

where $P_{\text{tr}}\left(N_w\right) = 1 - \left(1-\tau\right)^{N_w}$ with $P_{\text{tr}}\left(N_w\right)$ being the probability that there is at least one transmission in a time slot and τ being the transmission probability of each user. $P_{\text{s}}\left(N_w\right) = N_w \tau (1-\tau)^{N_w - 1} / P_{\text{tr}}\left(N_w\right)$, is the successful transmission probability, T_s is the average time that the channel is sensed busy because of a successful transmission, T_c is the average time that the channel is sensed busy by each station during a collision, T_σ is the duration of an unoccupied slot time, and $E\left[A\right]$ is the average packet size. In our model, the WiFi network adopts conventional distributed coordination function access and RTS/CTS access mechanisms. T_{c} and T_{s} are computed as done in [10].

We assume that one LTE time slot consists of T_W WiFi time slots. Based on the duty cycle mechanism, the UAVs can occupy ϑ fraction of T_W time slot on the unlicensed band while the WiFi users can occupy a fraction $(1-\vartheta)$ fraction of T_W time slots. Thus, the per WiFi user rate is:

$$R_w = \frac{R\left(N_w\right)\left(1-\vartheta\right)}{N_w}, \tag{2}$$

where N_w is the number of WiFi users on the unlicensed band. Given the rate requirement of each WiFi user γ, the fraction of the time slot on the unlicensed band allocated to the LTE-U users can be given by $\vartheta \le 1 - N_w \gamma / R(N_w)$.

B. UAV data rate analysis

Next, we define the rate of each user associated with a UAV. The UAVs' content transmission link consists of the wireless fronthaul links that connect each UAV to the cloud (ground-to-air links) and the UAV-users links (air-to-ground links). We consider probabilistic LoS and non-line-of-sight (NLoS) links over the licensed band for both the UAVs' fronthaul links and UAV-user links. In such a model, NLoS links experience higher attenuation than LoS links due to the shadowing and diffraction loss.

1) UAVs-users links over the licensed band: The LoS and NLoS path loss of UAV k transmitting a content to user i will be given by (in dB) [4]:

$$l_{ki}^{\text{LoS}} = 20\log\left(\frac{4\pi d_{ki} f}{c}\right) + \eta_{\text{LoS}}^{l}, l_{ki}^{\text{NLoS}} = 20\log\left(\frac{4\pi d_{ki} f}{c}\right) + \eta_{\text{NLoS}}^{l},$$

where $20\log\left(d_{ki} f 4\pi / c\right)$ is the free space path loss with d_{ki} being the distance between user i and UAV k, f being the carrier frequency, and c being the speed of light. η_{LoS}^{l} and η_{NLoS}^{l} represent, respectively, additional attenuation factors due to the LoS/NLoS connections over the licensed band. In our model, the probability of LoS connection depends on the environment, density and height of buildings, the locations of the user and the UAV, and the elevation angle between the user and the UAV. The LoS probability is given by [4]:

$$\Pr\left(l_{ki}^{\text{LoS}}\right) = \left(1 + X\exp\left(-Y\left[\phi_{ki} - X\right]\right)\right)^{-1}, \tag{3}$$

where X and Y are constants which depend on the environment (rural, urban, dense urban, or others) and $\phi_{ki} = \sin^{-1}\left(h_k / d_{ki}\right)$ is the elevation angle. Clearly, the average path loss from UAV k to user i is given by [4]:

$$\bar{l}_{ki}^{l} = \Pr\left(l_{ki}^{\text{LoS}}\right) \times l_{ki}^{\text{LoS}} + \Pr\left(l_{ki}^{\text{NLoS}}\right) \times l_{ki}^{\text{NLoS}}, \tag{4}$$

where $\Pr\left(l_{ki}^{\text{NLoS}}\right) = 1 - \Pr\left(l_{ki}^{\text{LoS}}\right)$. Based on the path loss, the downlink rate of user i associated with UAV k on the licensed band at time t can be given by:

$$R_{lki}(u_{ki}\left(t\right)) = u_{ki}\left(t\right) F_l \log_2\left(1 + \frac{P_K 10^{\bar{l}_{ki}^{l}/10}}{\sum\limits_{j \in \mathcal{K}, j \ne k} P_K 10^{\bar{l}_{ji}^{l}/10} + P_C h_i + \sigma^2}\right), \tag{5}$$

where F_l is the downlink bandwidth on the licensed band, P_K is the transmit power of each UAV, h_i is the channel gain between user i and the cloud, and P_C is the transmit power of the cloud over the fronthaul. σ^2 is the power of the Gaussian noise. Finally, $u_{ki}\left(t\right)$ is the fraction of the downlink licensed band allocated from UAV k to user i at time t with $\sum_i u_{ki}\left(t\right) = 1$.

2) UAVs-users links over the unlicensed band: In our model, the UAVs can only use the unlicensed band of the WiFi networks whenever the WiFi users' rate requirement is satisfied. Based on (2), we obtain a fraction T_V of a time slot over the unlicensed band that can be occupied by UAVs. Therefore, the downlink rate of user i associated with UAV k on the unlicensed band is:

$$R_{uki}(e_{ki}(t)) = e_{ki}(t)\vartheta F_u \log_2\left(1+\frac{P_K 10^{\bar{l}_{ki}^u/10}}{\sum\limits_{j\in\mathcal{K},j\neq k} P_K 10^{\bar{l}_{ji}^u/10}+\sigma^2}\right), \quad (6)$$

where $\bar{l}_{ki}^u$ is the average path loss over the unlicensed band, F_u is the bandwidth of the unlicensed band, and $e_{ki}(t)$ is the fraction of ϑ over the unlicensed band with $\sum_i e_{ki}(t) = 1$.

3) Cloud-UAVs ground-to-air links: The LoS and NLoS path loss from the cloud to UAV k can be given by [3]:

$$L_k^{\mathrm{LoS}} = d_{Ck}^{-\beta}, \qquad L_k^{\mathrm{NLoS}} = \varsigma d_{Ck}^{-\beta}, \quad (7)$$

where ς is the additional path loss of the NLoS connection and d_{Ck} is the distance between UAV k and the cloud. The average path loss $\bar{L}_k$ of the fronthaul link of UAV k can be computed using (3) and (4). Here, we assume that the total bandwidth of the UAVs' fronthaul is F_C which is equally divided among the users that received the contents from the cloud. Therefore, the fronthaul rate of each user associated with UAV k is given by:

$$R_{Ck}(t) = \frac{F_C}{U_C(t)}\log_2\left(1+\frac{P_C\bar{L}_k}{\sum\limits_{j\in\mathcal{K},j\neq k} P_K 10^{\bar{l}_{ki}^l/10}+\sigma^2}\right), \quad (8)$$

where $U_C(t)$ is the number of the users that receive a content from the cloud at time t. $U_C(t)$ can be calculated by the content server as the users request contents from the content server.

C. Queueing model

Let $A_i(t)$ be the random content arrival (number of bits) for user i from the content server at the end of time slot t. We assume that each user can request at most one content during each time slot t and, consequently, $A_i(t) \in \{0, L\}$. Let $Q_i(t)$ be the queue length (number of bits) of user i at the beginning of time slot t, which can be given by [13]:

$$Q_i(t+1) = Q_i(t) - R_{ki}(t) + A_i(t), \quad (9)$$

where $R_{ki}(t)$ is the rate of user i. Since the content transmission links consist of (a) UAV-user on the licensed band, (b) UAV-user on the unlicensed, (c) cloud-UAV-user on the unlicensed band, and (d) cloud-UAV-user on the licensed band, the rate of content transmission from UAV k to user i can be given by:

$$R_{ki}(u_{ki}(t), e_{ki}(t)) = \begin{cases} R_{lki}(u_{ki}(t)), & \text{link (a)}, \\ R_{uki}(e_{ki}(t)), & \text{link (b)}, \\ \frac{R_{uki}(e_{ki}(t))R_{Ck}(t)}{R_{uki}(e_{ki}(t))+R_{Ck}(t)}, & \text{link (c)}, \\ \frac{R_{lki}(u_{ki}(t))R_{Ck}(t)}{R_{lki}(u_{ki}(t))+R_{Ck}(t)}, & \text{link (d)}. \end{cases} \quad (10)$$

where the equation of link (c) is obtained from the fact that the time duration of a single data packet transmitted from the cloud to UAV k is $1/R_{Ck}(t)$ and a single data packet transmitted from UAV k to user i is $1/R_{uki}(t)$. Therefore, the data rate of the transmission from the cloud to user i is $\frac{1}{1/R_{Ck}(t)+1/R_{uki}(t)}$.

From (10), we can see that the rate of user i that receives a content from the UAV cache (link (a) and link (b)) is larger than the rate of user i that receives contents from the cloud (link (c) and link (d)). We use the notion of *queue stability* to measure the users' content transmission delay. In essence, a queue $Q_i(t)$ is said to be *rate stable* if [13]:

$$\lim_{t\to\infty}\frac{Q_i(t)}{t} = 0. \quad (11)$$

From [13, Theorem 2.8], we can also see that the queue $Q_i(t)$ is rate stable if $R_{ki}(t) \geq A_i(t)$.

D. Problem formulation

Given this system model, our goal is to develop an effective spectrum allocation scheme for cache-enabled UAVs that can allocate appropriate bandwidth over the licensed and unlicensed bands to satisfy the queue stability requirement of each user. To achieve this goal, we formulate an optimization problem whose objective is to maximize the number of users with stable queue. This maximization problem involves finding the optimal association $\mathcal{U}_k$ for each UAV k, bandwidth allocation indicators on the licensed band u_{ki}, time slots indicators on the unlicensed band e_{ki}, and the set of cached contents $\mathcal{C}_k$ for each UAV k. Therefore, this problem can be formalized as follows:

$$\max_{\boldsymbol{u},\boldsymbol{e},\mathcal{C}_k,\mathcal{U}_k} \sum_{k\in\mathcal{K}}\sum_{i\in\mathcal{U}_k} \mathbb{1}_{\left\{\lim\limits_{t\to\infty}\frac{Q_i(t)}{t}=0\right\}} = \sum_{k\in\mathcal{K}}\sum_{i\in\mathcal{U}_k} \mathbb{1}_{\{R_{ki}(u_{ki}(t),e_{ki}(t))\geq A_i(t)\}} \quad (12)$$

$$\text{s. t.} \quad R_w \geq \gamma, \quad (12a)$$

$$\sum_{i\in\mathcal{U}} u_{ki}(t) \leq 1, \quad \forall k \in \mathcal{K}, \quad (12c)$$

$$\sum_{i\in\mathcal{U}} e_{ki}(t) \leq 1, \quad \forall k \in \mathcal{K}, \quad (12d)$$

where $\mathbb{1}_{\{x\}} = 1$ when x is true and $\mathbb{1}_{\{x\}} = 0$ otherwise, $\mathcal{U}_k$ is the set of the users associated with UAV k, and $\boldsymbol{u}, \boldsymbol{e}$ denote the spectrum allocation indicators on the downlink licensed and unlicensed bands, respectively. (12a) guarantees the communication quality of each WiFi user. The LTE-U users can only occupy the unlicensed band when (12a) is satisfied. (12b) indicates that each user can only access each UAV's licensed band or unlicensed band, (12c) indicates that the licensed band allocation cannot exceed the total bandwidth for each UAV and (12d) captures the fact that the time slots over the unlicensed band cannot exceed the total number of time slots allocated to the UAVs.

III. Liquid State Networks for Content prediction and Self-Organizing Resource Allocation

The optimization problem in (12) is challenging to solve, because spectrum allocation and content caching depend on the user association which, in turn, depends on the rate of each user. In fact, this problem can be shown to be combinatorial and non-convex, thus it is difficult to solve it using conventional optimization algorithms. Moreover, each UAV may not know the users' content requests which makes it challenging to determine which content to cache at the UAVs. To address these challenges, we propose a novel *liquid state machine learning* approach [11] to predict the users' content request distribution and perform resource allocation.

Liquid state machine is a novel kind of spiking neural networks [11] that are randomly generated. Learning algorithms based on LSM can store the users' behavioral information and track the state of a network over time. Therefore, an LSM-based algorithm will enable the cloud to leverage information

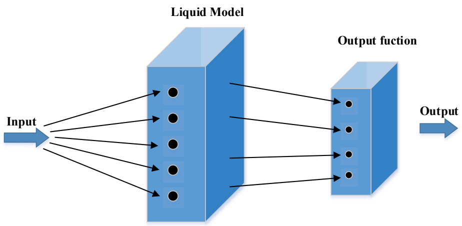

Fig. 1. The components of the proposed LSM-based algorithm.

on the users' behavior, that are stored in LSM, to predict the content request distribution and automatically adapt spectrum allocation to the change of network states. Consequently, LSM-based algorithms are promising candidates for content request distribution prediction and wireless resource allocation problems.

Next, we first begin by introducing the components of the LSM algorithm. Then, we introduce the entire process using LSM to predict the users' content request distribution and to solve (12).

A. LSM Components

As illustrated in Fig. 1, an LSM-based algorithm consists of five components: a) agents, b) input, c) output, d) liquid model, and e) output function. Since the prediction of the content request distribution and resource allocation are function-specific, we design the specific components for the problems of predicting content request distribution and resource allocation, separately.

1) Content request distribution prediction: The content request distribution prediction algorithm has the following components:

• *Agent*: The agent is the cloud. Since each LSM approach performs a content request distribution prediction for just one user, the cloud must implement U LSM algorithms.

• *Input:* The input of the LSM prediction algorithm is defined by a vector $\boldsymbol{x}_j\left(t\right)=\left[x_{j1}\left(t\right),\cdots,x_{jN_x}\left(t\right)\right]^{\mathrm{T}}$ that captures the *context information* related to user j's content request at time t. Such information includes age, gender, occupation, and device type (e.g., tablet or smartphone). N_x is the number of properties that constitute the context information of user j. The vector $\boldsymbol{x}_{t,j}$ is used to predict the content request distribution $\boldsymbol{y}_{t,j}$ of user j.

• *Output:* The LSM output at time t is a vector of probabilities $\boldsymbol{y}_j\left(t\right)=\left[p_{tj1},p_{tj2},\ldots,p_{tjN}\right]$ that represents the discrete probability density function of content request of user j with p_{tjn} being the probability that user j requests content n at time t.

• *Liquid model:* A liquid model for each user j can store the users' dynamic features that are extracted from the users' context over time. These dynamic features can be used with the output functions to predict the users' behavior such as content request and mobility patterns. Here, the liquid model consists of $W_1\times W_2\times W_3$ leaky integrate and fire neurons that are arranged in a 3D-column. In particular, each neuron that consists of resting state S, action state, and refractory period T_f (ms) is defined as a excitatory neuron while the neuron that consists of resting state, inhibitory state, and refractory period is defined as an inhibitory neuron. The resting state indicates that the neuron does not receive any users' context information. The action state indicates that the neuron receives a certain amount of users' context and transmits this context to the connected neurons. The inhibitory state indicates that the users' information in the neuron is decreasing and, thus, cannot reach an action state. When each neuron transmits information to other connected neurons, its users' information may decrease to a value that is below that of the resting state. In this case, the neuron will have a refractory period during which the neuron returns to the resting state. The action state of neuron j, $v_j\left(t\right)$ at time t is [11]:

$$v_j\left(t\right)=v_j\left(t-1\right)+\frac{S+ZI_j\left(t-1\right)-v_j\left(t-1\right)}{Z\rho},\tag{13}$$

where Z is the neuron resistance, $I_j\left(t-1\right)$ is the input of the users' information, and ρ is the neuron time constant. Based on (13), the LSM state can be given by $\boldsymbol{v}\left(t\right)=\left[v_1\left(t\right),v_2\left(t\right),\ldots,v_{N_W}\left(t\right)\right]$, where $N_W=W_1\times W_2\times W_3$ is the number of neurons.

The connections from the input to the liquid model are made with probability $\mathbb{P}_{\mathrm{IN}}$. The probability connection between neurons i and j can be given by:

$$\mathbb{P}_{ij}=Ce^{-\left(d\left(i,j\right)/\lambda\right)^2},\tag{14}$$

where $C\in\left\{C_{\mathrm{EE}},C_{\mathrm{EI}},C_{\mathrm{IE}},C_{\mathrm{II}}\right\}$ is a constant that depends on the type of both neurons. In particular, C_{EE} denotes an excitatory-excitatory connection, C_{EI} is an excitatory-inhibitory connection, C_{IE} is an inhibitory-excitatory connection, and C_{II} is a inhibitory-inhibitory connection. $d\left(i,j\right)$ is the Euclidean distance between neurons i and j. λ is a parameter that influences how often neurons are connected.

• *Output function:* The output function is used to build the relationship between the state of the LSM model and the prediction of each user's content request distribution. Let $\boldsymbol{f}_j\in\mathbb{R}^{N_W\times N}$ be the output function of UAV j, where N is the total number of contents. In order to predict the user's content request distribution, $\boldsymbol{f}_j$ is trained in an offline manner using ridge regression to approximate the prediction function:

$$\boldsymbol{f}_j=\boldsymbol{y}_{T,j}\boldsymbol{v}_j^{\mathrm{T}}\left(\boldsymbol{v}_j^{\mathrm{T}}\boldsymbol{v}_j+\delta^2\boldsymbol{I}\right)^{-1},\tag{15}$$

where $\boldsymbol{v}_j=\left[\boldsymbol{v}_j\left(1\right),\ldots,\boldsymbol{v}_j\left(N_t\right)\right]$ is the LSM state sequence for user j and N_M is the number of the prediction patterns of each user's content request distribution. Here, $\boldsymbol{y}_{T,j}$ is the target output of the LSM algorithm, $\boldsymbol{I}$ is an identity matrix, and δ is the learning rate. Based on the trained output function, the prediction of user j's content request distribution at time t can be given by:

$$\boldsymbol{y}_j\left(t\right)=\boldsymbol{f}_j\boldsymbol{v}_j\left(t\right).\tag{16}$$

2) Resource allocation: The LSM reinforcement learning algorithm for resource allocation has the following components:

• *Agent*: The agents in this LSM algorithm are the UAVs.

• *Input:* $\boldsymbol{m}_k(t)=\left[a_1(t),\cdots,a_{k-1}\left(t\right),a_{k+1}\left(t\right),a_K\left(t\right)\right]^{\mathrm{T}}$ represents the actions that all UAVs other than UAV k takes at time t. Here, $a_i\left(t\right)$ is the action that UAV i takes at time t. In particular, each UAV's action represents a UAV-user association scheme. As the UAV-user association is determined, the UAV cached content can be determined by our result in [7, Theorem 2]. Since the user association and cached contents [14] are determined, (12) for each UAV can be simplified as follows:

$$\begin{aligned}&\max_{\boldsymbol{u},\boldsymbol{e}}\sum_{i\in\mathcal{U}_{lkc}}\mathbb{1}_{\left\{R_{lki}\left(u_{ki}(t)\right)=A_i(t)\right\}}+\sum_{i\in\mathcal{U}_{ukc}}\mathbb{1}_{\left\{R_{uki}\left(e_{ki}(t)\right)=A_i(t)\right\}}+\\&\sum_{i\in\mathcal{U}_{lk}}\mathbb{1}_{\left\{\frac{R_{lki}\left(u_{ki}(t)\right)R_{Ck}(t)}{R_{lki}\left(u_{ki}(t)\right)+R_{Ck}(t)}=A_i(t)\right\}}+\sum_{i\in\mathcal{U}_{uk}}\mathbb{1}_{\left\{\frac{R_{uki}\left(e_{ki}(t)\right)R_{Ck}(t)}{R_{uki}\left(e_{ki}(t)\right)+R_{Ck}(t)}=A_i(t)\right\}},\end{aligned}\tag{17}$$

where $\mathcal{U}_{lkc}$ ($\mathcal{U}_{lk}$) is the set of users that are associated with UAV k over the licensed band and their requested contents are (not) stored in the cache. $\mathcal{U}_{ukc}$ ($\mathcal{U}_{uk}$) is the set of users that are associated with UAV k over the unlicensed band and their requested contents are (not) stored in the cache. (17) is a convex optimization problem that can be solved by linear programming.

- *Output:* The LSM output at time t is a vector of $\boldsymbol{b}_k(t) = [b_{k1}(t), b_{k2}(t), \ldots, b_{kA_k}(t)]$ that represents the resource allocation results. A_k is the number of actions that each UAV k can take and $b_{kj}(t)$ is the expected number of stable queue users when UAV k uses resource allocation scheme j, which is:

$$b_{ki}(t) = \sum_{\boldsymbol{a}_{-k} \in \mathcal{A}_{-k}} b_{ki,\boldsymbol{a}_{-k}}(a_{ki}, \boldsymbol{a}_{-k}) \pi_{-k,\boldsymbol{a}_{-k}}, \tag{18}$$

where $\mathcal{A}_{-k}$ is the action set of all UAVs other than UAV k and $b_{ki,\boldsymbol{a}_{-k}}$ is the number of stable queue users as UAV k uses the resource allocation scheme i and the other UAVs use the schemes $\boldsymbol{a}_{-k}$. $\pi_{-k,\boldsymbol{a}_{-k}} = \sum_{\boldsymbol{a}_{ki} \in \mathcal{A}_k} \pi(\boldsymbol{a}_{ki}, \boldsymbol{a}_{-k})$ is the marginal probability distribution over the action set of UAV k.

- *Liquid model:* The liquid model in the resource allocation algorithm is used to store the network state information including UAV-user association schemes and their corresponding output results. The liquid model consists of $W_1^\alpha \times W_2^\alpha \times W_3^\alpha$ number of leaky integrate and fire neurons that are arranged in a 3D-column. The generation of the resource allocation LSM model is similar to the one in the content request distribution prediction case.

- *Output function:* The output function is used to build the relationship between the UAVs-user association schemes and the number of users with a stable queue. Let $\boldsymbol{f}_k^\alpha \in \mathbb{R}^{A_k \times (A_k + N_W^\alpha)}$ be the output function of UAV k, where N_W^α is the number of the neurons in the liquid model. To train $\boldsymbol{f}_k^\alpha$, a linear gradient descent approach can be used to derive the following update rule,

$$\boldsymbol{f}_{k,i}^\alpha(t+1) = \boldsymbol{f}_{k,i}^\alpha(t) + \delta^\alpha (e_{k,i}(t) - b_{ki}(t)) [\boldsymbol{v}_k(t); \boldsymbol{m}_k(t)]^{\mathrm{T}}, \tag{19}$$

where $\boldsymbol{f}_{k,i}$ is row i of $\boldsymbol{f}_k$, δ^α is the learning rate, $e_{k,i}(t)$ is the expected output, and $\boldsymbol{v}_k(t)$ is the LSM state at time t. Based on (19), the estimated number of users that have a stable queue resulting from the users allocation scheme i is:

$$b_{ki}(t) = \boldsymbol{f}_{k,i}^\alpha [\boldsymbol{v}_k(t); \boldsymbol{m}_k(t)]. \tag{20}$$

B. LSM Algorithm for content prediction and spectrum allocation

To solve the problem in (12), the cloud first predicts the content request distribution of each user using an LSM-based prediction approach. Based on the users' content request distribution, each UAV k uses the LSM-based learning algorithm with ϵ-greedy mechanism [10] to find the optimal users association. Once the users association is determined, the optimal content caching and spectrum allocation will also be determined. In this algorithm, each UAV can store the users' and network's states as UAV k adopts different users association schemes. During each iteration, the LSM algorithm can record the number of stable queue users, $b_{ki,\boldsymbol{a}_{-k}}(a_{ki}, \boldsymbol{a}_{-k})$. Since the LSM-based algorithm satisfies the convergence conditions of [10, Theorem 2], as time elapses, each UAV k's output resulting from the resource allocation scheme i will converge to a final value b_{ki}. At this convergence point, b_{ki} indicates the expected value of the number of stable queue users with respect to all other UAVs' stategies. The proposed LSM approach performed by the cloud and each UAV k is shown, in detail, in Algorithm 1.

Algorithm 1 LSM-based learning algorithm

Input: The set of users' context, $\boldsymbol{x}_j(t)$, UAVs' input $\boldsymbol{m}_k(t)$;
Init: The cloud generates the liquid model for each user.
Each UAV generates a liquid model based on (13) and (14).
Calculate the time slots L based on (2)
Predict users' content request distribution using (16)
for time t **do**
  Estimate the number of the users that are at stable queuing state using (20)
  if $rand(.) < \varepsilon$ **then**
    Randomly choose one action
  else
    Choose action $a_k(t) = \arg\max_{a_k(t)} (\boldsymbol{b}_k(t))$
  end if
  Observe the number of the users that are at the stable queuing state $e_{k,i}(t)$
  Update the output weight matrix $\boldsymbol{f}_{k,i}^\alpha(t)$ based on (19)
  Update the input $\boldsymbol{a}_k(t)$ according to the result of the users choosing BSs
end for

TABLE I
SYSTEM PARAMETERS

Parameter	Value	Parameter	Value	Parameter	Value
P_K	15 dBm	L	1 Mbits	Z	20 dB
P_C	20 dBm	$\eta_{\mathrm{LoS}}^u, \eta_{\mathrm{NLoS}}^u$	1.2, 23	ρ	30 ms
σ	-94 dBm	DIFS	50 μs	F_C	2 Gbit
F_l	10 Mbit	$E[A]$	1500 bytes	N_w	8
F_u	20 Mbit	δ, δ^α	0.1, 0.05	β	2
C	3	W_1, W_2, W_3	5,5,20	ACK	304 μs
CTS	304 μs	X, Y	11.9, 0.13	RTS	352 μs
N	25	$\eta_{\mathrm{LoS}}^l, \eta_{\mathrm{NLoS}}^l$	1, 20	ς	20 dB
C_{EE}	0.3	C_{IE},C_{II},C_{EI}	0.2,0.1,0.4	SIFS	16 μs
R_w	4 Mbps	$W_1^\alpha, W_2^\alpha, W_3^\alpha$	5,5,30	$\mathbb{P}_{\mathrm{IN}}$	0.3
W	2	S	13.5 mV	N_x	4

IV. Simulation Results

In our simulations, the content request data that the LSM uses to train and predict content request distribution is obtained from *Youku* of *China network video index*[1]. The detailed parameters are listed in Table I. We consider a circular cloud-based UAVs network area with a radius $r = 200$ m, $U = 20$ uniformly distributed users and $K = 5$ uniformly distributed UAVs. For implementing the proposed LSM-based algorithm, we use the Matlab LSM toolbox described in [11]. Other system parameters are listed in Table I. We compare our approach with: a) Q-learning algorithm in [10] with content caching and b) Q-learning algorithm without content caching. All statistical results are averaged over 5000 independent runs.

In Fig. 2, we show how the average number of stable queue users changes as the number of the UAVs varies. From Fig. 2, we can see that the number of stable queue users increases as the number of the UAVs increases. This is due to the fact that increasing the number of UAVs provides more connection options for the users, and, thus, improves the number of stable queue users. Fig. 2 also shows that the proposed LSM algorithm can yield up to 33.3% and 50.3% gains in terms of the number of stable queue users compared to Q-learning algorithm with cache and Q-learning without cache, respectively, for a network with 5 UAVs. These gains stem from the fact that the proposed LSM algorithm can use the historical resource allocation information to find an optimal resource allocation scheme and predict the users' content request distribution to improve content caching.

In Fig. 3, we show the variations of two content request probabilities of an arbitrarily selected user during one day. From Fig. 3, we can see that the accuracy of the predictions of the proposed LSM algorithm is within less than 8% from the real content request probability. Since this gap does not affect the ranking of each content request probability, the cloud can find the optimal contents to cache using the proposed algorithm. Fig. 3 also shows that the sum of the probabilities with which this user requests contents 1 and 2 exceeds 0.5 during each hour. This is because the user always requests a small number of contents

Fig. 4 shows the number of iterations needed till convergence

[1]The data is available at http://index.youku.com/.

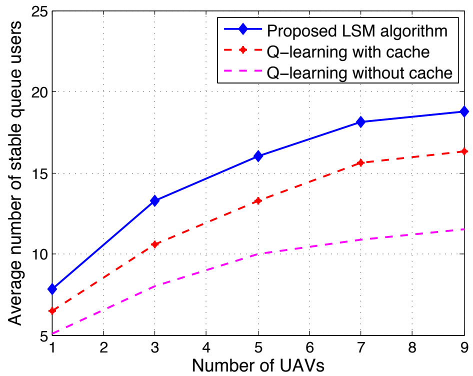

Fig. 2. Average number of stable queue users as the number of UAVs varies.

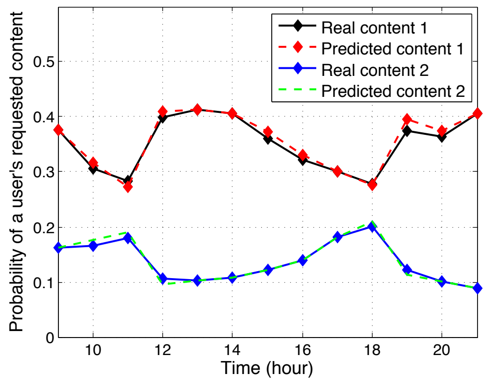

Fig. 3. Content request probability predictions.

for both the proposed approach and Q-learning with cache. In this figure, we can see that, as time elapses, the number of stable queue users increases until convergence to their final values. Fig. 4 also shows that the proposed approach needs 400 iterations to reach convergence and exhibits a considerable reduction of 33.3% less iterations compared to Q-learning with cache. This is due to the fact that LSM algorithm stores the users states.

In Fig. 5, we show the variations of the number of the users that are allocated to the licensed and unlicensed bands as the number of contents stored at the UAV cache varies. From Fig. 5, we can see that the number of users over the licensed and unlicensed bands increases as the number of cached contents increases. This is due to the fact that content caching decreases the traffic load of the cloud-UAV links thus decreasing the rate needed for having stable queues at the users. Fig. 5 also shows that the number of users on the licensed band is 50% more than the number of users over the unlicensed band due to the difference in the parameters of the path loss over the unlicensed band.

V. Conclusion

In this paper, we have developed a novel framework that uses flying cache-enabled UAVs to provide service for users in an LTE-U system. We have formulated an optimization problem that seeks to maximize the number of stable queue users. To solve this problem, we have developed a novel algorithm based on the machine learning tools of liquid state network. The proposed prediction algorithm enables the cloud to predict each user's content request distribution and, thus, determine the UAV's cached contents. Using the proposed LSM resource allocation algorithm, each UAV can decide on its spectrum allocation scheme autonomously with limited information on the network state. Simulation results have shown that the proposed approach yields significant performance gains.

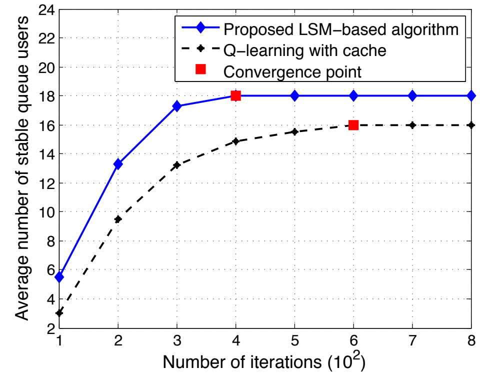

Fig. 4. Convergence of the learning algorithms.

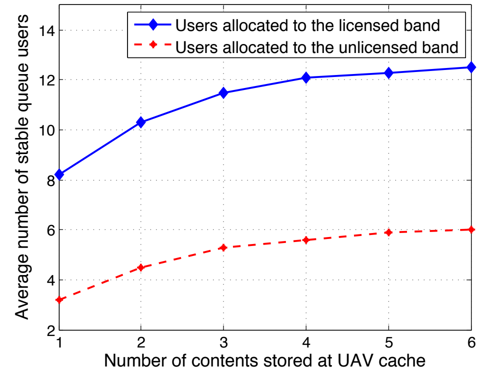

Fig. 5. Number of users associated with licensed and unlicensed bands as the number of stored contents at UAV cache varies.

References

[1] Y. Zeng, R. Zhang, and J. L. Teng, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, 2016.
[2] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage," *IEEE Communications Letters*, vol. 20, no. 8, pp. 1647–1650, Aug. 2016.
[3] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 3949–3963, June. 2016.
[4] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," in *Proc. of IEEE Global Communications Conference*, Austin, TX, USA, Dec. 2014.
[5] I. Bor-Yaliniz and H. Yanikomeroglu, "The new frontier in RAN heterogeneity: Multi-tier drone-cells," *IEEE Communications Magazine*, Nov. 2016.
[6] T. X. Tran and D. Pompili, "Octopus: A cooperative hierarchical caching strategy for cloud radio access networks," in *Pro. of International Conference on Mobile Ad Hoc and Sensor Systems*, Brasilia, Brazil, Oct. 2016.
[7] M. Chen, M. Mozaffari, W. Saad, C. Yin, M. Debbah, and C. S. Hong, "Caching in the sky: Proactive deployment of cache-enabled unmanned aerial vehicles for optimized quality-of-experience," *IEEE Journal on Selected Areas on Communications (JSAC), Special Issue on Human-In-The-Loop Mobile Networks*, vol. 35, no. 5, pp. 1046–1061, May 2017.
[8] D. Athukoralage, I. Guvenc, W. Saad, and M. Bennis, "Regret based learning for UAV assisted LTE-U/WiFi public safety networks," in *Proc. of IEEE Global Communications Conference (GLOBECOM), Mobile and Wireless Networks Symposium,*, Washington, DC, USA, December 2016.
[9] U. Challita, L. Dong, and W. Saad, "Deep learning for proactive resource allocation in lte-u networks," in *Pro. of European Wireless Conference*, Dresden, Germany, May 2017.
[10] M. Chen, W. Saad, and C. Yin, "Echo state networks for self-organizing resource allocation in LTE-U with uplink-downlink decoupling," *IEEE Transactions on Wireless Communications*, vol. 16, no. 1, pp. 3–16, Jan. 2016.
[11] W. Maass, "Liquid state machines: Motivation, theory, and applications," *Computability in context: Computation and logic in the real world*, pp. 275–296, 2010.
[12] Y. Pan, C. Pan, H. Zhu, Q. Z. Ahmed, M. Chen, and J. Wang, "On consideration of content preference and sharing willingness in D2D assisted offloading," *IEEE J. Select. Areas Commun. (JSAC), Special Issue on Human-In-The-Loop Mobile Networks*, vol. 35, no. 4, Apr. 2017.
[13] M. J. Neely, "Stochastic network optimization with application to communication and queueing systems," *Synthesis Lectures on Communication Networks*, vol. 3, no. 1, pp. 1–211, 2010.
[14] D. Wang, P. Ren, Q. Du, L. Sun, and Y. Wang, "Security provisioning for miso vehicular relay networks via cooperative jamming and signal superposition," *IEEE Transactions on Vehicular Technology*, to appear, 2017.