Autonomous Robustness Control for Fog Reinforcement in Dynamic Wireless Networks

Beatriz Lorenzo, *Senior Member, IEEE*, Francisco Javier González-Castaño, Linke Guo, *Senior Member, IEEE*, Felipe Gil-Castiñeira, and Yuguang Fang, *Fellow, IEEE*

Abstract— The sixth-generation (6G) of wireless communications systems will significantly rely on fog/edge network architectures for service provisioning. To realize this vision, AI-based fog/edge enabled reinforcement solutions are needed to serve highly stringent applications using dynamically varying resources. In this paper, we propose a cognitive dynamic fog/edge network where primary nodes (PNs) temporarily share their resources and act as fog nodes (FNs) for secondary nodes (SNs). Under this architecture, that unleashes multiple access opportunities, we design distributed fog probing schemes for SNs to search for available connections to access neighbouring FNs. Since the availability of these connections varies in time, we develop strategies to enhance the robustness to the uncertain availability of channels and fog nodes, and reinforce the connections with the FNs. A robustness control optimization is formulated with the aim to maximize the expected total long-term reliability of SNs' transmissions. The problem is solved by an online robustness control (ORC) algorithm that involves online fog probing and an index-based connectivity activation policy derived from restless multi-armed bandits (RMABs) model. Simulation results show that our ORC scheme significantly improves the network robustness, the connectivity reliability and the number of completed transmissions. In addition, by activating the connections with higher indexes, the total long-term reliability optimization problem is solved with low complexity.

Index Terms— Fog/edge, latency, reliability, restless multi-armed bandit (RMAB), robustness.

I. INTRODUCTION

FUTURE 6G wireless systems will be expected to fulfill new communications, networking and computing requirements for a massive connectivity, mostly expecting low-latency

and high-reliability, in a highly dynamic scenario [1], [2]. Fog/edge computing, which enables computing anywhere along the cloud-to-thing continuum, is seen as the way forward to meet these stringent requirements [3], [4]. In [31], fog learning is proposed to distribute machine learning (ML) model training from devices to cloud servers and compensate the limitations of centralized ML for battery-limited devices, and latency-sensitive and privacy sensitive applications. These use cases need reliable communication links that have low probability of failure to achieve low latency. Many works [32] have analyzed the performance in terms of reliability but without specific solutions to improve it. Akbar *et al.* [6] estimate the reliability level of links using a k-nearest neighbor algorithm and an adaptive decision mechanism to select the best path for different types of applications. However, the dynamic nature of fog/edge architectures with services hosted within smartphones, gateways, and user-provided access points results in uncertain link availability, which must be considered when selecting a reliable path. Furthermore, with the rapidly increasing connectivity demands, spectrum resources are becoming scarce and thus, solutions to enhance the reliability must account for spectrum availability as well.

The incorporation of cognitive radio capabilities into fog/edge architectures has attracted much attention recently due to its ability to increase spectrum efficiency and throughput by unleashing multiple access opportunities at the network edge [3], [7]–[8]. In our previous works [3], [5], [7] we presented a self-organized data and spectrum trading algorithm to harvest available resources at the network edge, improving significantly the revenue of the operator. Si *et al.* [8] studied proactive caching of popular video contents over harvested bands to maximize the spectrum utilization. By a proper design, the multiple access opportunities can be used to increase the robustness of the network defined as the probability that the network remains connected under traffic dynamics. Achieving a high network robustness is crucial to reconfigure the connections on time by jointly allocating available channels and fog nodes to meet high reliability and low latency requirements at the network edge.

Resource allocation and service provisioning under fog/edge architectures have been studied in several works [9]–[12]. Gu *et al.* [9] investigated the joint radio and computational resource allocation to satisfy user QoS and proposed a matching game framework to solve the problem. Wang *et al.* [10] studied joint computation offloading and content caching as a convex problem, and solved it with the alternating direction method of multipliers (ADMM). These problems are solved

either in static environments or by adapting static algorithms to the network dynamics and developing heuristic algorithms. Recently, ML has emerged as a powerful tool to make fog/edge computing highly adaptable and enable fast-reconfiguration [13]–[16]. Abdulkareem *et al.* [13] addressed the use of ML for autonomous intelligent management and operation in fog-aided IoT. Chen *et al.* [14] studied proactive network association and anticipatory mobility management through ML. In [15], a traffic-flow prediction algorithm based on a long short-term memory (LSTM) was presented to predict and control the mobile-traffic flow of the entire network. Sun *et al.* [16] presented a task offloading algorithm for vehicular edge computing systems based on Multi-Armed Bandits (MAB) that enabled vehicles to learn the offloading delay performance of their neighboring vehicles. However, none of the above papers provided solutions to improve the connection reliability in cognitive fog/edge networks given the network dynamics nor to reduce the impact of the latter on the latency. Predicting the connectivity availability will reduce the number of reconfigurations needed and improve the resource utilization since fewer resources will be wasted due to link failures. Besides, we can reinforce the connections by keeping channels and fog nodes with high availability probability as backups to meet the most stringent requirements.

In this paper, we contribute to the reinforcement of fog/edge networks by presenting a robustness control framework to access dynamically varying resources at the network edge using fog nodes, and serve applications with low-latency and high-reliability requirements. In particular, the main contributions of this work are the following:

a) We design distributed fog probing schemes to search for the availability of mobile edge devices to act as fog nodes, and the availability of spectrum and computing resources. First, an agnostic fog probing scheme is developed that assumes no prior knowledge on the outage probability of the connection, which may result into temporal recapturing of the channels used by SNs and/or fog nodes due to PNs' traffic. The process is modelled as a two dimensional absorbing Markov chain. The probability of successful transmission, which depends on the activity of PNs and fog nodes, is quantified. This scheme is used as a benchmark for comparison with the autonomous schemes developed later.

b) Fog reinforcement strategies are presented to enhance the robustness to the uncertain availability of channels and fog nodes by probing multiple connections with high probability of success as backups. A comprehensive framework to model and analyze these strategies is elaborated. The robustness optimization problem is formulated as a stochastic optimization problem. The aim is to maximize the expected long-term network performance in terms of reliability defined as the probability of transmitting successfully within a latency bound. However, solving this optimization is complex since it is a combinatorial problem and its complexity increases exponentially with the network size.

c) Restless Multi-Armed Bandits (RMAB) provide an efficient way to derive an index-based robustness control algorithm with low computational complexity. However, investigating the problem structure to cast it as an RMAB is challenging [28]. By reducing our original Markov connectiv-ity state model, we reformulate the robustness control problem in the form of a RMAB, which enables an autonomous implementation. Without loss of generality, the reduced model allows direct implementation of the Whittle index policy with significantly low complexity. The problem is solved by an online robustness control (ORC) algorithm that integrates online fog probing and index-based connectivity activation policy. We prove the indexability of the connectivity activation policy theoretically and obtain the Whittle index in closed-form. By relaxing the constraint of the number of connections activated per slot, the Whittle index is the optimal solution to our RMAB.

d) We evaluate our algorithms through numerous simulations. First, the evaluation is conducted in small-sized networks to compare the performance with the original problem formulation that requires high computational time for large-sized networks. Then we evaluate the index-based scheduling algorithm for large-sized networks. We compare the performance with some typical scheduling algorithms. The results demonstrate that our approach significantly outperforms existing solutions.

The remainder of this paper is organized as follows. Section II describes the related work. The system model is elaborated in Section III. In Section IV, the fog reinforcement framework and the agnostic fog probing scheme are described. The robustness control optimization is formulated in Section V and it is solved in Section VI using our proposed ORC algorithm. Section VII evaluates the proposed algorithms through simulations. Concluding remarks are provided in Section VIII.

II. RELATED WORK

Several works have investigated computing resource failures in cloud and fog computing services. Yao and Ansari [17] studied the reliability of virtual machines as their probability of failures in fog nodes when processing computing tasks. They formulated a multi-objective optimization problem to assign computing tasks to virtual machines (VMs) in a fog node. Dantu *et al.* [18] utilized smartphones as fog nodes and designed a software architecture to provide reliability and adaptability. Yao and Ansari [19] addressed the joint optimization of power control and fog resource provisioning in terms of the number of VMs to guarantee task completion time requirements. Liao *et al.* [34] propose a scheme to balance computing resources in edge IoT by monitoring computing demand.

Some works have studied channel assignments in cognitive networks through spectrum leasing [35], channel switching and rerouting [36]–[37] but without reliability and latency guarantees and with fixed access points. Recently, a few works have addressed fog resource provisioning under network dynamics in fog networks. Zhao *et al.* [20] presented a multi-tier operations scheduler to optimize node assignments at the control tier and resource allocation at the access tier under dynamic constraints. They solved the problem using Lyapunov optimization techniques and developed an online scheduling algorithm that achieved at least half of the optimal value. Omoniwa *et al.* [21] utilized fog nodes as relays and proposed a relay scheme to minimize the transmission

TABLE I

NOTATION

i, j, b	Index of SN, PN, and channel
N, M, B	Total number of SNs, PNs, and channels
$\mathcal{N}, \mathcal{M}, \mathcal{B}$	Set of SNs, PNs, and channels
a_{ij}^b	Availability probability of channel b at link $i \to j$
α_j	Availability probability of PFN j
l_{ij}^b	Link availability probability between i and j on channel b
C_{ij}	Capacity of link $i \to j$
$\tau_{ij}^{b,a}, \tau_{ij}^t, \tau_{ij}^c, \tau_{ij}^b$	Access, transmission, and computing delay, and overall
$Q_i, Q_{r,i}$	Total and remaining data required by task of SN i
x_{ij}^b, y_{ij}^b	Indicator connectivity constraint, and association between
ρ^r	Robustness of the network under reinforcement strategy r
O_{ij}^b	Outage in the transmission between i and j on channel b
$\mathcal{P}_i(t)$	Fog probing strategies
ξ_j	Reliability of the transmission of SN i
n_c, n_a	Number of backup channels and backup fog nodes
$R_i(t), R_i'(t)$	Reward of SN from activation and from probing
c_k, c_k'	Cost of transmission failure and of probing
ω_k	Belief of connection k
S_j, S_b, S_k, s_i	State of PFN j, channel b, connection k, and SN i
$v(s)$	Whittle Index in state s

outage that jointly optimizes mobility and power consumption. Yang *et al.* [22] considered fog nodes equipped with cognitive capabilities and studied the joint optimization of spectrum utilization and energy efficiency in collaborative task offloading. In our previous work [5], we presented a Robust Dynamic Network Architecture (RDNA) together with a holistic cross-layer approach to improve network robustness. In this paper, our focus is on optimizing robustness under resource uncertainty to meet latency and reliability requirements. A comprehensive analytical framework is developed to model and analyze robustness enhancement, which encompasses fog probing and fog reinforcement strategies.

III. SYSTEM MODEL

In this section, we describe our proposed network architecture, and the related communication and computing models. The most important notations used in the paper are summarized in Table I.

A. Network Architecture

We consider a cognitive, dynamic fog/edge network architecture, as illustrated in Fig. 1. A set of primary nodes (PNs) $\mathcal{M} = \{1, 2, \ldots, M\}$, such as smartphones, tablets, etc., with data storage, computation and packet forwarding capabilities share their connectivities and act as primary fog nodes (PFNs). The PFNs serve a set of secondary nodes (SNs) $\mathcal{N} = \{1, 2, \ldots, N\}$ with limited capabilities, such as smart sensor nodes that collect data for machines, objects, etc. The PFNs will collect the data from SNs, perform necessary computation and distribute it throughout the network. We assume that SNs are equipped with cognitive capabilities to harvest available frequency channels in the set $\mathcal{B} = \{1, 2, \ldots, B\}$. The network is operated by a primary operator (PO) that incentivizes its users, whenever their terminals are idle, to act as PFNs. The high density of user terminals provides many connectivity alternatives and opportunities to establish backup
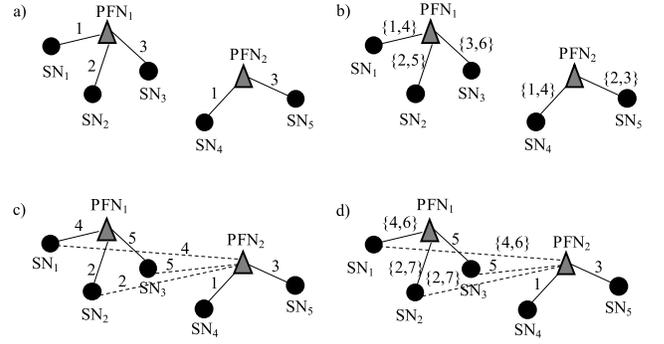
Fig. 1. Illustration of fog reinforcement (the selected connections have the highest success probability as explained in Section V): a) conventional topology without reinforcement; b) topology with backup channels; c) topology with backup PFNs; and d) topology with backup channels and PFNs.

connections to improve the robustness of the network against traffic dynamics.

B. Communication Model

We assume that SNs are equipped with one radio that can be tuned into any available frequency channel for message delivery. The availability of frequency channels varies in time and space as these channels may be occupied by PNs' transmissions. Thus, PNs activity will affect the performance of SNs' transmissions. We assume slotted transmissions in the primary and secondary networks with different arrival/departure times in each network. Let $a_{ij}^b, i \in \mathcal{N}, j \in \mathcal{M}$, denote the probability that channel b at link $i \to j$ is available for SN i transmission to PFN j and $(1 - a_{ij}^b)$ the probability that channel b at link $i \to j$ is occupied by a PN transmission, and thus, unavailable for SN i transmission. In addition, we denote the availability of PFN j, $j \in \mathcal{M}$ to act as an access point by α_j. PFN j is available to act as an access point and share its resources when it is not transmitting/processing its own traffic.

To characterize the transmission/interference in the physical layer, we adopt the widely accepted protocol model [23]. The power propagation gain from SN i to PFN j is $g_{ij} = \beta \cdot d_{ij}^{-\chi}$, where β is an antenna-related parameter, χ is the path loss factor, and d_{ij} is the distance between the two nodes. According to the Shannon-Hartley theorem, if SN i, $i \in \mathcal{N}$ transmits data to PFN j, $j \in \mathcal{M}$, using available channel b, the link capacity will be

$$C_{ij} = W \log_2 \left(1 + P_i \cdot g_{ij}/\gamma\right) \tag{1}$$

where $W^b = W$ is the bandwidth of channel b, P_i is the transmission power at SN i and γ is the Gaussian noise power at PFN j. In the following section, we will address the transmission constraints to avoid interference.

If Q_i is the amount of data required for the task of SN i, the transmission time needed to transmit its traffic is $\tau_{ij}^t = Q_i/C_{ij}$ and the energy consumption is $e_{ij}^t = P_i \tau_{ij}^t$.

C. Computation Model

Suppose f_j is the computation capacity of PFN j in CPU cycles per unit of time. Denote by w_j its current workload – the portion of processing capacity currently occupied by PFN

j's own tasks. Then the available processing capacity to share with SNs is $\Gamma_j = f_j(1 - w_j)$. In addition, we denote by δ_i the amount of computing resource required by the task of SN i in CPU cycles. Accordingly, the execution time for computing the task of SN i at PFN j is $\tau_{ij}^c = \delta_i/\Gamma_j$, and the energy consumption at PFN j is $e_{ij}^c = \varepsilon_j \tau_{ij}^c$, where ε_j is the energy cost per CPU cycle [24]. Given the limitation of mobile terminals in computational capacity, we assume that at most one task is executed at a time.

IV. FOG REINFORCEMENT FRAMEWORK

By taking advantage of the high density of user terminals, in the subsequent development we present the methods to identify possible backup connectivities (and their constraints) that could be used to reinforce the connections of SNs with the fog under uncertainty of available channels and PFNs.

A. Interference Constraints and Connectivity Availability

We consider scheduling of SNs transmissions in the frequency domain, i.e., channel assignments for transmission and receptions to ensure that there is no interference at the same node and among adjacent nodes.

Denote by $\mathcal{B}_i \subseteq \mathcal{B}$ the set of available channels at SN $i \in \mathcal{N}$. Suppose that channel b is available at SN i and PFN j, i.e., $\mathcal{B}_{ij} = \mathcal{B}_i \cap \mathcal{B}_j$. Define

$$x_{ij}^b = \begin{cases} 1, \text{if SN } i \text{ can transmit data to PFN } j \text{ on channel } b \\ 0, \text{otherwise} \end{cases} \quad (2)$$

For an SN $i \in \mathcal{N}$ and a channel $b \in \mathcal{B}$, the set of PFNs that can use channel b and are within the transmission range R_i^T of SN i is $\mathcal{T}_i^b = \{j | d_{ij} \leq R_i^T, j \neq i, b \in \mathcal{B}_{ij}\}$. Note that a PFN $j \in \mathcal{M}$ cannot receive from multiple SNs on the same channel,

$$\sum_{\{i | j \in \mathcal{T}_i^b\}} x_{ij}^b \leq 1. \quad (3)$$

Likewise, if SN i uses channel b for transmitting data to PFN $j \in \mathcal{T}_i^b$, then any other SN that can interfere with PFN j should not use this channel,

$$x_{ij}^b + \sum_{\{m \in \mathcal{T}_n^b\}} x_{nm}^b \leq 1, n \in \mathcal{I}_j^b, n \neq i \quad (4)$$

where $\mathcal{I}_j^b = \{n | d_{nj} \leq R_j^I, n \neq j, b \in \mathcal{B}_{nj}, \mathcal{T}_n^b \neq \emptyset\}$ is the set of SNs that can interfere with the reception of PFN j on channel b, $\mathcal{B}_{nj}$ is the set of the licensed channels available to SN n and PFN j, and $\mathcal{T}_n^b \neq \emptyset$ indicates that SN n has a PFN to which it can transmit by interfering with reception at PFN j.

A feasible scheduling of SNs transmissions in frequency channels must satisfy the previous interference constraints. The scheduling of PNs transmissions is out of the scope of this paper. Nevertheless, since the availability of channels for SNs transmissions and the availability of PFNs depend on the traffic in the primary network, we model PNs' traffic as well for completeness of the model.

Denote the set of PNs whose transmission on channel b will interrupt SN i transmission to PFN j as $\mathcal{C}_j^b = \{z | d_{zj} \leq R_j^I, z \neq j, b \in \mathcal{B}_{zj}, \mathcal{T}_z^b \neq \emptyset\}$ where channel b is available at
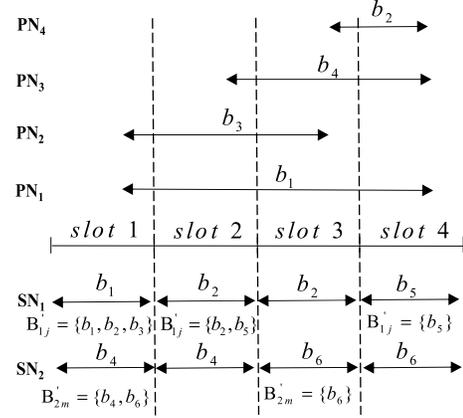
Fig. 2. Illustration of the scheduling process after a PN return.

PN z and PFN j. By slightly abusing the notation $\mathcal{T}_z^b \neq \emptyset$ indicates that z has a destination to which it can transmit by interrupting the transmission between SN i and PFN j.

At the beginning of slot t, each SN probes the availability of channels and fog nodes that satisfy its connectivity requirements. Let us elaborate the availability probability of channel b for SN transmission. For analytical tractability of the model, we assume that traffic arrivals follow a Poisson process [3]. The probability of having Z arrivals from PNs' in the set $\mathcal{C}_j^b$ within a time slot of duration Δt is $p_Z(\Delta t) = e^{-\lambda_{\mathcal{C}_j^b} \Delta t} \cdot (\lambda_{\mathcal{C}_j^b} \Delta t)^Z / Z!$ where $\lambda_{\mathcal{C}_j^b} = \sum_{z \in \mathcal{C}_j^b} \lambda_z$. Then, the availability of channel b for SN i transmission at slot t is obtained as the probability that channel $b \in \mathcal{B}_{ij}$ has not been recently allocated (i.e., in the previous Δt) to any arrivals from PN $z \in \mathcal{C}_j^b$,

$$a_{ij}^b(t) = 1 - \sum_{Z=1}^{\infty} \frac{Z}{|\mathcal{B}_{ij}|} p_Z(\Delta t^-) \quad (5)$$

where Δt^- refers to the previous Δt and $Z/|\mathcal{B}_{ij}|$ is the probability that Z arrivals are allocated to a particular channel (i.e., channel b) out of $|\mathcal{B}_{ij}|$. Here, we assume PNs can access any channel with the same probability. Equation (5) can be modified to capture other PN channel access policies. The scheduling process after a PN return is illustrated in Fig. 2. At the beginning of slot 1, SN$_1$ has a set $\mathcal{B}'_{1j}$ of available channels to transmit to PFN j (found by probing), and selects to transmit in channel b_1. After the transmission is initiated, PN$_3$ returns to channel b_1 interrupting the transmission. At the beginning of the next slot, SN$_1$ finds the new set of available channels and transmits in channel b_2. It continues transmitting in this channel until the transmission is interrupted by PN$_4$ in slot 3. Finally, in slot 4 SN$_1$ finds the new set of channels and transmits in channel b_5. A similar behavior is followed by SN2. For simplicity, we illustrate only the impact of PNs' activity on channel availability but its extension to the fog availability is straightforward.

A PN j is available as a PFN and performs computing tasks for SNs if it is not transmitting/processing its own traffic. We denote by $p_J(\Delta t) = e^{-\lambda_{P_j} \Delta t} (\lambda_{P_j} \Delta t)^J / J!$ the probability that PN j receives J new requests from its own traffic with arrival rate λ_{P_j} within Δt. Thus, the probability that PN j is available to serve as a fog node in slot t is the probability of

not receiving any new request ($J = 0$) in Δt^-

$$\alpha_j(t) = p_{J=0}(\Delta t^-) \quad (6)$$

Since the availability of fog nodes (and channels) changes in time, a SN may transmit to different fog nodes and channels in subsequent time slots. If a SN transmission is interrupted, it will be repeated it in the next slot using the same or different fog nodes depending on the availability.

Hence, the probability that the link between SN i and PFN j on channel b is available at time t is

$$l_{ij}^b(t) = a_{ij}^b(t)\alpha_j(t) \quad (7)$$

The connection will be successful if the link remains available for the entire duration of the slot Δt. Denote the outage O_{ij}^b as the probability that the connection is interrupted due to either a return of a PN to the currently allocated channel b (as illustrated in Fig. 2) and/or a new request from PFN j's own traffic in Δt. Thus, the transmission between SN i and PFN j on channel b at time t will be successful with probability

$$l_{ij}^b(t + \Delta t) = l_{ij}^b(t)(1 - O_{ij}^b(\Delta t)) \quad (8)$$

B. Reinforcement Strategies

Strategies to reinforce the connectivity with the fog and increase the robustness of the network connections will be presented. Robustness is the property of the end devices to remain connected to the fog and providing service under dynamically varying traffic. Dynamic traffic induces uncertainty to the availability of connectivity (i.e., channels and PFNs) which impacts on latency and reliability, as well as on overall network performance.

Definition 1: The latency τ refers to the time elapsed since the data is transmitted until it is received by the destination. The latency τ_{ij}^b between i and j on channel b includes the access delay $\tau_{ij}^{b,a}$ of SN i to PFN j and it will be elaborated in Section V depending on the reinforcement strategy used, the transmission time τ_{ij}^t, and the computational time τ_{ij}^c. Downlink time is negligible compared to uplink data offloading time and computation, hence, it has not been considered in the calculus [25],

$$\tau_{ij}^b = \tau_{ij}^{b,a} + \tau_{ij}^t + \tau_{ij}^c \quad (9)$$

Definition 2: Reliability ξ refers to the probability of successful transmission within a latency bound τ_{max}. Therefore, the reliability of the connection of SN i is

$$\xi_i = \Pr\left(\sum_j \sum_b \tau_{ij}^b \leq \tau_{max,i}\right) \quad (10)$$

where τ_{ij}^b is the latency.

In the following, we present our strategies to reinforce the connectivity with the fog. We represent a reinforcement strategy as the pair (n_c, n_a) where the first element indicates the number of backup channels and the second one is the number of backup fog nodes. The selection of the specific backup connections is explained in Section V. We define the following four strategies: $r = 0 \rightarrow (0,0)$, $r = 1 \rightarrow (n_c, 0)$, $r = 2 \rightarrow (0, n_a)$, $r = 3 \rightarrow (n_c, n_a)$ that indicate no

reinforcement, reinforcement through n_c backup channels, reinforcement through n_a backup PFNs, and reinforcement through both n_c backup channels and n_a backup fog nodes, respectively. For simplicity in the sequel, we remove the time dependency t in the link availability probability (7).

a) No reinforcement ($r = 0$): It describes the conventional connectivity option where an SN $i \in \mathcal{N}$ probes the availability of a PFN $j \in \mathcal{T}_i^b$ on one channel b at a time. This is illustrated in Fig. 1a. The probability that the link between i and j is available on channel b under this strategy is obtained by (7) as

$$l_{ij}^{b,r=0} = a_{ij}^b\alpha_j \quad (11a)$$

b) Backup channels ($r = 1$): An SN $i \in \mathcal{N}$ probes a set of backup channels $\mathcal{B}'_{ij} \subset \mathcal{B}_{ij}$ to increase the probability that there is a link available for transmission to a PFN $j \in \mathcal{T}_i^b$. This is illustrated in Fig. 1b. The probability that the link between i and j is available either on channel b or on any backup channel $k \in \mathcal{B}'_{ij}$ out of $n_c = |\mathcal{B}'_{ij}|$ backup channels is

$$l_{ij}^{b \cup \mathcal{B}'_{ij}, r=1} = l_{ij}^{b,r=0} + (1 - a_{ij}^b)\sum_{k=1,k\in\mathcal{B}'_{ij}}^{n_c} a_{ij}^k\alpha_j$$
$$\times \prod_{s=1}^{k-1}(1 - a_{ij}^s) \quad (11b)$$

The first term is the link availability probability between i and j on channel b with no reinforcement as in expression (11a). The second term indicates the probability that channel b is not available and the probability that the link between i and j is available on a backup channel k where s is the index of the backup trial.

c) Backup PFNs ($r = 2$): An SN $i \in \mathcal{N}$ probes a set of PFNs on channel b denoted by $(\mathcal{T}_i^b)' \subset \mathcal{T}_i^b$, as shown in Fig. 1c. By introducing $n_a = |(\mathcal{T}_i^b)'|$ backup PFNs, the link availability probability between i and $j \cup (\mathcal{T}_i^b)'$ on channel b is

$$l_{ij\cup(\mathcal{T}_i^b)'}^{b,r=2} = l_{ij}^{b,r=0} + (1 - \alpha_j)\sum_{m=1,m\in(\mathcal{T}_i^b)'}^{n_a} a_{im}^b\alpha_m$$
$$\times \prod_{q=1}^{m-1}(1 - \alpha_q) \quad (11c)$$

The first term denotes the link availability probability with no reinforcement as expressed in (11a), and the second term is the probability that PFN j is not available and the probability that there is a backup PFN m available to serve SN i on channel b. Index q indicates the backup trial.

d) 2-level backup ($r = 3$): This strategy combines backup channels and fog nodes as illustrated in Fig. 1d. Thus, an SN $i \in \mathcal{N}$ probes $1 + n_c$ channels to transmit to any of its $1 + n_a$ PFNs. The probability that there is a link available between i and $j \cup (\mathcal{T}_i^b)'$ on channel $b \cup \mathcal{B}'_{ij}$ under this strategy is given in (11d). The first term is the probability of link availability under reinforcement strategy $r = 1$ given in (11b). The second term is the probability of link availability under reinforcement strategy $r = 3$ as expressed in (11c). The third term is the link availability under no reinforcement as in (11a). Finally, the fourth term is the probability that the channel b and PFN j are not available for SN i's transmission multiplied by the probability that SN i has a connection available on a backup channel k to a backup PFN m. The indexes s and q are the indexes of the channel and PFN backup trails, respectively.
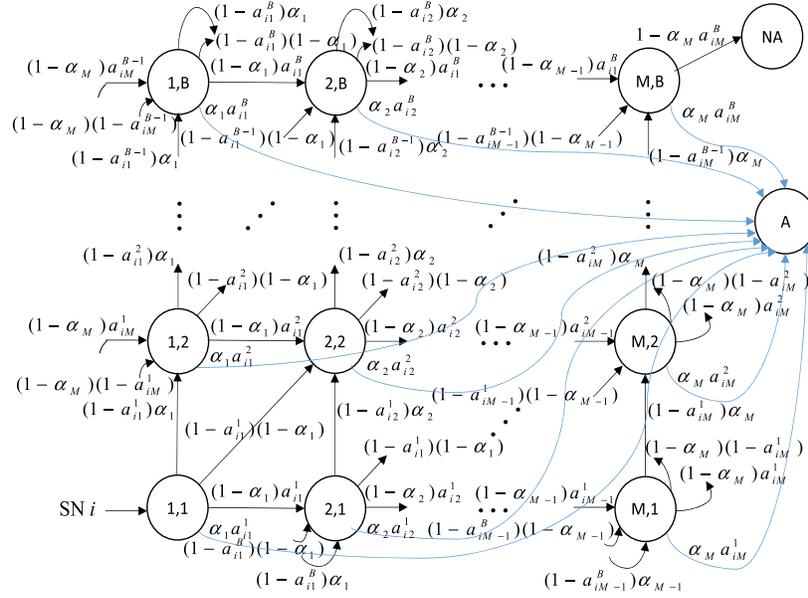
Fig. 3. Fog access state transition diagram with $S_a = (j_a, b_a)$.

$$l_{ij\cup(\mathcal{T}_i^b)'}^{b\cup\mathcal{B}_{ij}',r=3} = l_{ij}^{b\cup\mathcal{B}_{ij}',r=1} + l_{ij\cup(\mathcal{T}_i^b)'}^{b,r=2} - l_{ij}^{b,r=0}$$

$$+(1-a_{ij}^b)(1-\alpha_j)\sum_{k=1,k\in\mathcal{B}_{ij}'}^{n_c}\sum_{m=1,m\in(\mathcal{T}_i^b)'}^{n_a}$$

$$\times\alpha_m\prod_{q=1}^{m-1}(1-\alpha_q)a_{im}^k\prod_{s=1}^{k-1}(1-a_{iq}^s) \tag{11d}$$

For simplicity, we have limited our previous discussion to obtaining the link availability probability $l_{ij}^{b,r}(t)$ under different reinforcement strategies. Its extension to obtaining the probability of successful transmission $l_{ij}^{b,r}(t+\Delta t)$ for each strategy r is straightforward from (8).

Definition 3: The robustness of the network is defined as the probability that the network remains connected under dynamically varying traffic, which is given as the average probability that users have at least one connection available to transmit reliably,

$$\rho^r = \sum_i \rho_i^r/N = \sum_i \left(1-\prod_j\prod_b\left(1-l_{ij}^{b,r}(t+\Delta t)\right)\right)/N \tag{12}$$

where r is the reinforcement strategy used, $j \in \mathcal{T}_i^b$, and $l_{ij}^{b,r}(t+\Delta t)$ is the probability of successful connectivity (8) under strategy r. Section V describes how SNs learn to probe the set of connections that have high probability of successful transmission.

C. Agnostic Fog Probing

We model SNs access to the fog network by using a fog probing mechanism used to check the availability of channels and adjacent fog nodes at the beginning of each slot. Initially, we assume SNs do not have any preliminary knowledge about whether the probed connections are successful. This scheme is referred to as *agnostic* fog probing, which will be used as a benchmark for comparison with fog probing schemes that incorporate learning in Section V. The probing process is represented by an absorbing Markov chain with *MB* transient states, where M is the number of fog nodes and B the number of channels, and the two absorbing states "A" and "NA" indicating access and non-access to the fog, respectively. The fog network access state transition diagram is shown in Fig. 3 where each transition state is given by the pair (j, b) denoting the indexes of the PFN and the channel, respectively.

The agnostic fog probing protocol works as follows. At the beginning of a slot, each SN i checks the availability of its adjacent PFNs, starting with its most preferable one, and the availability of each channel randomly. Let us assume that SN i starts checking the availability in state $(1,1)$. If PFN $j = 1$ is not available, the process moves rightward with probability $(1 - \alpha_1)a_{i1}^1$, whereas if channel $b = 1$ is not available, the process moves upward with probability $(1 - a_{i1}^1)\alpha_1$. Similarly, if neither PFN $j = 1$ nor channel $b = 1$ are available, the process will move up along the diagonal with probability $(1 - a_{i1}^1)(1 - \alpha_1)$. If a PFN and a channel are available in state $S_a = (j_a, b_a)$, the process will move to the "A" state with probability $a_{i1}^1\alpha_1$. Based on the reinforcement strategy, this process may be repeated, starting from the next state S_a+1 until the SN finds all available PFNs and channels. The process will finish when all PFNs and channels are checked and there are no more available, finishing in the "NA" state. If a PN returns to a channel currently allocated to a SN or the allocated PFN receives a new task from its own traffic, the connection will be interrupted, and the process will be repeated to find a new connection at the beginning of the next slot, as shown in Fig. 3.

To obtain the access delay $\tau_{ij}^{b;a}$ under the agnostic fog probing scheme, we define the transition probability matrix of fog network access states $\mathbf{S} = \|S(j, b; j', b')\| = \|S(m, m')\|$ with indexes $m = j + M(b-1)$ and $m' = j' + M(b'-1)$, $m, m' = 1, 2, \ldots, MB$ denoting the current and next state, respectively. Following the theory of absorbing Markov chains [26], [27], we arrange matrix $\mathbf{S}$ in a canonical form as

$$\mathbf{S}^{re} = \left\| \begin{matrix} \mathbf{I} & \mathbf{0} \\ \mathbf{R} & \mathbf{S} \end{matrix} \right\| \tag{13}$$

with size $N_S \times N_S$, where N_S is the number of states (i.e, $N_S = N_A + MB$), $\mathbf{I}$ is a $N_A \times N_A$ unity matrix corresponding to N_A absorbing states, $\mathbf{0}$ is a $N_A \times MB$ all-zero matrix, $\mathbf{R}$ is an $MB \times N_A$ matrix of transition probabilities from transient states to absorbing states, and $\mathbf{S}$ is an $MB \times MB$ matrix of transition probabilities between transient states. These matrices are outlined in the Appendix.

We define the fundamental matrix as $\mathbf{N} = (\mathbf{I} - \mathbf{S})^{-1}$ with dimensions $MB \times MB$. The mean access time for the process to reach an absorbing state starting from transient state m is the mth entry of the vector [26]

$$(\tau_1^a, \ldots, \tau_{MB}^a)^t = T\mathbf{N}\mathbf{1} \qquad (14)$$

when the dwell time for any state m is the same, $T = T_m$ and $\mathbf{1}$ is an $MB \times 1$ column vector of all ones. The generalization to any dwell time is straightforward. The variance of each τ_m^a can be expressed as [26]

$$\operatorname{var} \tau_m^a = 2\mathbf{N}\mathbf{T}\mathbf{S}\mathbf{N}\mathbf{e} + \mathbf{N}(\mathbf{e}_{sq}) - (\mathbf{N}\mathbf{e})_{sq}$$

where $\mathbf{T}$ is an $MB \times MB$ diagonal matrix with elements T_m, $\mathbf{e}$ is a column vector of the same elements, and $\| \, \|_{sq}$ is the square of each component of $\| \, \|$. The access delay $\tau_{ij}^{b,a}$ is obtained as in (14) with $T = 1$ and the inverse mapping $(j,b) \leftarrow m$.

The probability that transient state m is absorbed to the absorbing state $n = \{$"A", "NA"$\}$ is the (m,n)-entry of the matrix

$$\mathbf{B} = [b_{mn}] = \mathbf{N}\mathbf{R} \qquad (15)$$

which provides the probability of available transmission in state $(j,b) \leftarrow m$ or equivalently from SN i to PFN j on channel b.

In the agnostic fog probing, SNs will probe all connections until they find an available one. This process is costly since it consumes time and network resources, especially if connectivity reinforcement strategies are used. Therefore, as a next step we will incorporate learning in the fog probing process to reduce the number of probed connections to those with the highest probability of transmission success.

V. ROBUSTNESS CONTROL OPTIMIZATION

In this section, we formulate the robustness control optimization problem to maximize the long-term reliability of SNs' transmissions. The problem is solved in two steps. First, each SN i probes a set of connections $\mathcal{P}_i(t)$ at time t based on its reinforcement strategy. After the fog probing is completed, each SN notifies the secondary operator (SO) on the availability of the probed connections. Let us recall that SNs' transmissions use the channels and PFNs from the primary network whenever available. Once all notifications are received, the SO decides which connections to activate in each slot and receives a reward R_i for each successful connection of SN i, which depends on the reliability. The objective for the SO is to activate the connections to maximize the long-term total discounted reward.
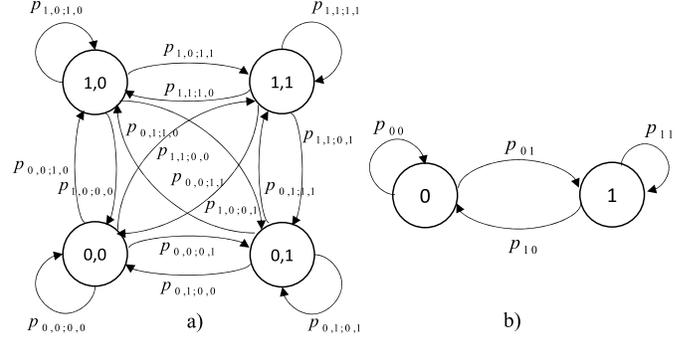
Fig. 4.   Markov connectivity state model: a) Original and b) Reduced.

A. Automatized Fog Probing and Reinforcement Selection Strategy

By using the agnostic fog probing scheme described in Section IV.C, SNs check the availability of the connections in the current instant but they are unaware of the probability of transmitting successfully (i.e., there is no outage in the transmission) by using these connections. To improve the selection of the connections, we present an automatized fog probing scheme based on RMABs model [28], [29] that builds a belief, based on previous experience, that the connection will be successful. In the automatized fog probing, at the beginning of time slot t, each SN i chooses a set $\mathcal{P}_i(t) \subset \{1, 2, \ldots, K\}$ of connections to probe ($1 \leq K < MB$) and receives a reward if the probed connections remain available. The objective is to probe the links with the highest probability of being available for the slot duration. We model the availability of each connection $(j,b) \rightarrow k \in \mathcal{P}_i(t)$ as a Markov process with four possible states, depending on the availability "1" or unavailability "0" of PFN j and channel b: $S_{j,b}(t) = (S_j(t), S_b(t)) = (0,0), (0,1), (1,0),$ or $(1,1)$. The first index indicates the state of PFN $S_j(t)$ and the second one the state of channel $S_b(t)$. The state of each connection evolves from slot to slot as a Markov chain with transition matrix $\mathbf{P}^{org} = [p_{S_{j,b};S'_{j,b}}]$ as illustrated in Fig. 4a.

For tractability of the model and to cast our problem as a RMAB and obtain the Whittle index, we elaborate an equivalent reduced Markov model, as shown in Fig. 4b, with two states, $S_k(t) = 1$ and $S_k(t) = 0$, depending on the availability or unavailability of the connection k, respectively. The state of the probed connection is now obtained as $S_k(t) = S_j(t) \cdot S_b(t)$, where $S_j(t)$ is the state of the PFN and $S_b(t)$ is the state of the channel. The transition matrix of the reduced Markov model is $\mathbf{P}^{red} = [p_{S_k S'_k}]$, where the transition probabilities of connection k during slot t are

$$p_{00}^k(t + \Delta t) = 1 - l_{ij}^b(t)(1 - O_{ij}^b(\Delta t)) \qquad (16)$$

$$p_{01}^k(t + \Delta t) = l_{ij}^b(t)(1 - O_{ij}^b(\Delta t)) \qquad (17)$$

$$p_{10}^k(t + \Delta t) = O_{ij}^b(\Delta t) \qquad (18)$$

$$p_{11}^k(t + \Delta t) = (1 - O_{ij}^b(\Delta t)) \qquad (19)$$

where k is the index of the connection of SN i to PFN j on channel b. The probability of connection outage O_{ij}^b is defined in Section IV.A. The transition probability $p_{00}^k(t + \Delta t)$ from $S_k(t) = 0$ to $S_{k'}(t + \Delta t) = 0$ is obtained as the probability that the connection remains unavailable. The connection transition

probability $p_{01}^k(t + \Delta t)$ from state $S_k(t) = 0$ to state $S_{k'}(t + \Delta t) = 1$ is equal to the probability of successful transmission. The connection transition probability $p_{10}^k(t + \Delta t)$ from state $S_k(t) = 1$ to state $S_{k'}(t + \Delta t) = 0$ is equal to the connection outage. Finally, the connection transition probability $p_{11}^k(t + \Delta t)$ from state $S_k(t) = 1$ to state $S_{k'}(t + \Delta t) = 1$ is equal to the probability of no connection outage.

Since the connection state $S_k(t)$ is not observable until when the availability of the connection is checked, we define a belief matrix $\mathbf{\Omega}(t) = [\omega_1(t), \ldots, \omega_K(t)]$, where $\omega_k(t)$ is the conditional probability that $S_k(t) = 1$. It has been shown that the conditional probability that each connection is in state 1 given all past decisions and observations is a sufficient statistic for optimal decision making [28]. Given the fog probing selection $\mathcal{P}_i(t)$ and the observation at time t, the belief state in time $t + 1$ can be obtained recursively as

$$\omega_k(t+1) = \begin{cases} p_{11}^{(k)}, & k \in \mathcal{P}(t), S_k(t) = 1 \\ p_{01}^{(k)}, & k \in \mathcal{P}(t), S_k(t) = 0 \\ \mathcal{T}(\omega_k(t)), & k \notin \mathcal{P}(t) \end{cases} \quad (20)$$

where $\mathcal{T}(\omega_k(t)) = \omega_k(t)p_{11}^{(k)} + (1 - \omega_k(t))p_{01}^{(k)}$ is the belief operator when connection k has not been probed in the current slot. If there is no prior information on the initial system state, the kth entry of the initial belief vector $\mathbf{\Omega}(1)$ can be set to the steady state probability that the reduced system is in state 1, $\pi_1^{(k)}$. To calculate $\pi_1^{(k)}$, next we analyze the relation between the steady-state probabilities of the original system, shown in Fig. 4a, and the reduced system, shown in Fig. 4b.

Denote by $\boldsymbol{\pi}^{org} = (\pi_{00}^{org}, \pi_{01}^{org}, \pi_{10}^{org}, \pi_{11}^{org})$ the steady state distribution for the Markov chain with transition matrix $\mathbf{P}^{org} = [p_{S_{j,b}; S'_{j,b}}]$ satisfying

$$\boldsymbol{\pi}^{org} = \boldsymbol{\pi}^{org}\mathbf{P}^{org} \quad (21)$$

and by $\boldsymbol{\pi}^{red} = (\pi_0^{red}, \pi_1^{red})$ the steady state distribution for the reduced Markov chain with transition matrix $\mathbf{P}^{red} = [p_{S_k S'_k}]$ satisfying

$$\boldsymbol{\pi}^{red} = \boldsymbol{\pi}^{red}\mathbf{P}^{red} \quad (22)$$

The equivalence between the original system and the reduced system is described as

$$\begin{aligned} \pi_1^{red} &= \pi_{11}^{org} \\ \pi_0^{red} &= \pi_{00}^{org} + \pi_{01}^{org} + \pi_{10}^{org} \end{aligned} \quad (23)$$

where $\pi_0^{red} + \pi_1^{red} = 1$ and $\pi_{00}^{org} + \pi_{01}^{org} + \pi_{10}^{org} + \pi_{11}^{org} = 1$. Since the values of the transition matrix $\mathbf{P}^{org}$ can be obtained from the traffic arrival and departure distributions, $\boldsymbol{\pi}^{org}$ is calculated by solving the system (21). Then, $\boldsymbol{\pi}^{red}$ is obtained by solving (23). Finally, by solving system (22) the steady-state probabilities of the reduced system are calculated

$$\pi_0^{(k)} = \frac{p_{10}^{(k)}}{p_{01}^{(k)} + p_{10}^{(k)}}; \quad \pi_1^{(k)} = \frac{p_{01}^{(k)}}{p_{01}^{(k)} + p_{10}^{(k)}}$$

By formulating the fog probing as an RMAB, each connection k represents an arm and the belief state $\omega_k(t)$ is the state of the arm at time t. Each SN will choose a number of arms K to probe at each slot while the other arms are unobserved.

The belief state shows the states of both probed and unprobed arms.

Each SN i receives a reward at time t when a connection that has probed successfully (i.e., available connection) is activated by the SO minus the cost of probing the connections. Based on the reinforcement strategy, each SN will probe $K = (1 + n_c)(1 + n_a)$ connections to increase the chances of finding one with a high belief and, thus, increase the probability that it will be selected by the SO and receive a reward,

$$R_i'(t) = \sum_{k \in \mathcal{P}_i(t)} y_k(t)\omega_k(t) - Kc_k' \quad (24)$$

where $y_k(t) \in \{0, 1\}$: $y_k(t) = 1$, if the connection has been selected by the SO, or $y_k(t) = 0$, otherwise, and c_k' is the cost of probing the connection. If SN i has a reinforcement strategy $r = 0$, then it will probe only $K = 1$ connections since the number of backup channels n_c and backup fog nodes n_a will be zero. In our fog probing problem, arms are stochastically identical (i.e., all arms have the same Markovian dynamics and reward structure). Thus, we focus on deriving the optimal policy at an individual slot since it would remain the same for different slots [28]. We define an *online* fog probing policy obtained by probing K arms with the highest belief in each slot. The online fog probing policy $\hat{\mathcal{P}}_i(t)$ is then given by

$$\hat{\mathcal{P}}_i(t) = \underset{\mathcal{P}_i(t)}{\arg\max} R_i'(t) \quad (25)$$

B. Connectivity Activation Policy

Once the fog probing is completed, each SN sends the information about the availability of the probed connections to the SO. Given that all SNs share any available channels and PFNs to transmit, the SO will activate the SNs' connections to maximize the reward subject to the interference constraints (3) and (4). The indicator $y_{ij}^b(t) = 1$ is used to denote that connection $(j, b) \rightarrow k \in \mathcal{P}_i(t)$ has been allocated for SN i transmission, and $y_{ij}^b(t) = 0$ otherwise. If the activated connection k results in a successful transmission the SO receives a reward r_k. Otherwise, it receives a penalty c_k.

The problem is to determine which connections to activate at each time slot to maximize the expected total discounted reward for the SO,

$$\begin{aligned} \max_{y_{ij}^b(t)} \quad & \mathbf{E}\left[\sum_{t=1}^{\infty} \sum_i \beta^{t-1} R_i(s_i(t), y_{ij}^b(t)) \right] \\ \text{subject to} \quad & \sum_i y_{ij}^b(t) \leq \gamma B \\ & y_{ij}^b(t) \in \{0, 1\}, (j, b) \rightarrow k \in \mathcal{P}_i(t) \end{aligned} \quad (26)$$

where β $(0 < \beta \leq 1)$ is the discount factor, $R_i(s_i(t), y_{ij}^b(t))$ is the reward of the SO that will be elaborated in the next section, and $y_{ij}^b(t) \in \{0, 1\}$ denotes the association between SN $i \in \mathcal{N}$ and PFN $j \in \mathcal{M}$ on channel $b \in \mathcal{B}$ for data transmission and task computation in time t. Since the available channels are shared among multiple SUs $(B < N)$, to avoid interference, the number of connections activated simultaneously are constrained to a fraction γ of the available channels B. The fraction γ can be obtained using an interference graph [30] and considering the interference constraints defined in (4) and (5) with $y_{ij}^b(t) \leq x_{ij}^b(t)$.

Solving the previous optimization is complex since it is a combinatorial problem [11]. To make it tractable, in the next section we formulate (26) as a RMAB problem and derive a connectivity activation index policy based on Whittle Index [28], [29], [38].

VI. INDEX-BASED CONNECTIVITY ACTIVATION POLICY

By modelling the SO connectivity activation policy as a RMAB problem, we view each connection as an arm. When an arm is activated, the corresponding SN can transmit. To formulate the RMAB, we define the decision epoch, state space, state transition probability, and reward.

1) Decision Epochs: Time is divided into discrete time slots and decisions are made each time t, $t \in \{1, 2, \ldots, \infty\}$.

2) State Space: Recall that each SN i wants to transmit an amount of data Q_i in $\tau_{max,i}$ slots. We define the state of the SN i transmission $s_i(t) = (Q_{r,i}(t), \tau_{r,i}(t))$, indicating that SN i has a remaining amount of data $Q_{r,i}(t)$ to transmit and $\tau_{r,i}(t)$ remaining slots to complete it. The system state at decision t consists of the states of all SNs $\mathbf{s} = (s_1(t), \ldots, s_N(t))$. At each time t, there are N SNs waiting to transmit in one of their available probed connections. Note that the state s_i of the SN i is different from the state of the connection S_k in the previous section.

3) Action: At each decision epoch, the action $y_{ij}^b(t)$ taken by the SO determines which SNs can transmit. If SO takes action $y_{ij}^b(t) = 1$ the connection $(j, b) \to k$ for SN i is activated, otherwise $y_{ij}^b(t) = 0$. Since there are $B < N$ channels, the action taken at any time t should satisfy the constraint $\sum_i y_{ij}^b(t) \leq \gamma B$.

4) State Transition Probability: For each SN i, the state $s_i(t)$ will transfer to different states with probability $\Pr\{s_i(t+1)|s_i(t), y_{ij}^b(t)\}$ depending on the action of the SO and the uncertain availability of the links. When $\tau_r = 1$, assuming that SN i have packets to transmit periodically, $\Pr\{s_i(t+1)|s_i(t), y_{ij}^b(t)\} = 1$ independently of the action $y_{ij}^b(t)$ taken and the state will transfer to $s_i(t+1) = (Q_i, \tau_{\max,i})$. Similarly, $\tau_r > 1$ and $Q_r = 0$, independently of the action we have $\Pr\{s_i(t+1)|s_i(t), y_{ij}^b(t)\} = 1$ with $s_i(t+1) = (0, \tau_r - 1)$. However, if $\tau_r > 1$ and $Q_r > 0$ and the connection is activated the state will transfer to $s_i(t+1) = (Q_{r,i}(t) - Q_i(t+1), \tau_{r,i}(t) - 1)$ with probability ω_{ij}^b or to state $s_i(t+1) = (Q_{r,i}(t), \tau_{r,i}(t) - 1)$ with probability $1 - \omega_{ij}^b$. The probability that the connection will be successful or belief ω_{ij}^b is obtained as in (20). If the connection is not activated, $\Pr\{s_i(t+1)|s_i(t), y_{ij}^b(t)\} = 1$ with $s_i(t+1) = (Q_{r,i}(t), \tau_{r,i}(t) - 1)$.

5) Reward: The reward of SN i at time t depends on the current state and the action taken,

$$R_i(s_i(t), y_{ij}^b(t))$$
$$= \begin{cases} 0, & Q_{r,i}(t) = 0 \\ y_{ij}^b(t)(\omega_{ij}^b(t)r_{ij}^b(t) - (1 - \omega_{ij}^b(t))c_{ij}^b(t)) \\ + (1 - y_{ij}^b(t))c_{ij}^b(t), & Q_{r,i}(t) > 0, \tau_{r,i}(t) > 0 \end{cases} \quad (27)$$

where $\omega_{ij}^b \leftarrow \omega_k$ is the belief state as in (20), $c_{ij}^b(t) \leftarrow c_k$ is the penalty incurred when the connection is not available or it has not been activated, and $r_{ij}^b(t) \leftarrow r_k$ is the reward of connection k obtained when the connection is successful.

To obtain the reward $r_{ij}^b(t)$ per slot t, we can rewrite the reliability, defined in (10), as

$$\xi_i = p(\tau_i \leq \tau_{\max,i}) = \sum_{t=1}^{\tau_{\max,i}} \sum_j \sum_b y_{ij}^b(t)\omega_{ij}^b(t)r_{ij}^b(t) \quad (28)$$

where $r_{ij}^b(t) = Q_{ij}^b(t)/Q_i$ is the ratio of the amount of data transmitted in time t with respect to the overall amount of data Q_i that SN i aims to transmit.

A. Indexability and Whittle Index Policy

Whittle index policy is the optimal solution to a Lagrangian relaxation of RMABs [28]. In our problem, this is achieved by relaxing the constraint in (26) in which the number of activated arms can vary over time given that their discounted average over the infinite horizon equals γB,

$$\mathbf{E}\left[\sum_{t=1}^{\infty} \sum_i \beta^{t-1} y_{ij}^b(t)\right] = \frac{\gamma B}{1 - \beta}$$

Based on the Lagrangian multiplier theorem, the RMAB can be decomposed into a single-arm activation problem and so, it suffices to consider a single arm (i.e., connection),

$$\max_y \mathbf{E}\left[\sum_{t=1}^{\infty} \sum_i \beta^{t-1}(R_i(s_i(t), y(t)) - \lambda y(t))\right]$$

The aim is to decide whether to activate the arm at each slot based on the concept of subsidy for passivity [28]. Let us construct a single-bandit process identical to the one previously described except for a constant subsidy ν that is obtained when the arm is passive. The ν-subsidy reward is formally given by

$$R_i^\nu(s_i(t), y(t)) = R_i(s_i(t), y(t)) + \nu\mathbf{1}(y(t) = 0) \quad (29)$$

where $\mathbf{1}(\cdot)$ equals 1 if the expression in the bracket is true, and 0 otherwise. The SO decides whether to activate an arm or not at each time t to maximize the total discounted ν-subsidy reward

$$V_i^\nu(s) = \sum_{t=1}^{\infty} \beta^{t-1} R_i^\nu(s_i(t), y(t)) \quad (30)$$

with initial state s. To simplify the notation, we drop the subscripts i and t without loss of generality.

Let $V(s)$ denote the value function that represents the maximum expected total discounted reward that can be accrued from a single-arm bandit process when the initial state is s and two actions, $y = 0$ and $y = 1$, are possible:

$$V(s) = \max\{V(s, y = 0), V(s, y = 1)\} \quad (31)$$

where $V(s; y)$ is the expected total discounted reward when action y is taken at the first slot followed by the optimal policy in future slots as

$$V(s, y = 0) = R(s, 0) + \nu + \sum_{s' \in S} \beta p(s'|s, 0)V^\nu(s') \quad (32)$$

$$V(s, y = 1) = R(s, 1) + \sum_{s' \in S} \beta p(s'|s, 1)V^\nu(s') \quad (33)$$

The term $p(s'|s, y)$ denotes the probability that SN i changes from state s to next state s' when decision y is taken and $V^\nu(s')$ is the total discounted future reward. In (32),

$V(s, y = 0)$ is given by the sum of the ν-subsidy reward in the first slot under action $y = 0$ and the total discounted future reward. Likewise, $V(s, y = 1)$ is obtained.

Definition 4: The Whittle index $\nu_i(s)$ of an arm i in state s is the infimum subsidy ν that makes the two decisions (activating arm i or not) equally rewarding:

$$\nu_i(s) = \inf_{\nu}\{\nu : V(s, y = 0) \geq V(s, y = 1)\} \quad (34)$$

where $V(s; y)$ is the expected reward when action y is taken at state s.

Definition 5: An arm is indexable if the passive set $\mathcal{Z}(\nu) = \{\nu : V(s, y = 0) \geq V(s, y = 1)\}$ of the single-armed bandit process with subsidy ν monotonically increases as ν increases from $-\infty$ to $+\infty$. An RMAB is indexable if every arm is indexable.

To establish the indexability and derive the closed-form expression of the Whittle index, we distinguish the following cases in calculating the expected total discounted reward:

1) When $\tau_r = 1$, assuming that SNs have packets to transmit periodically, $p(s'|s, y) = 1$ independently of the action y taken and the SN will change to $s' = (Q, \tau_{\max})$. The remaining time to complete the transmission is initialized to $\tau_{\max}$ and the SN can start a new transmission in the next slot.

- If $Q_r = 0$ and $y = 0$ we have $V((0, 1), 0) = \nu + \beta V(Q, \tau_{\max})$, whereas if $y = 1$ we obtain $V((0, 1), 1) = \beta V(Q, \tau_{\max})$. Therefore, by (34) the Whitte index is $\nu(0, 1) = 0$.

- If $Q_r > 0$ and $y = 0$, we have that $V((Q_r, 1), 0) = \nu - c + \beta V(Q, \tau_{\max})$, whereas if $y = 1$ we have $V((Q_r, 1), 1) = \omega r - (1 - \omega)c + \beta V(Q, \tau_{\max})$. Then, the Whittle index is $\nu(Q_r, 1) = \omega r + \omega c$.

The previous derivations establish the indexability of the problem when $\tau_r = 1$.

2) When $\tau_r > 1$,

- If $Q_r = 0$ independently of the action y taken, $p(s'|s, y) = 1$ with $s' = (0, \tau_r - 1)$. If $y = 0$, $V((0, \tau_r), 0) = \nu + \beta V(0, \tau_r - 1)$ is obtained, whereas if $y = 1$, we have $V((0, \tau_r), 1) = \beta V(0, \tau_r - 1)$ and the Whittle index is $\nu(0, \tau_r) = 0$.

- If $Q_r > 0$ and $y = 0$, we have $p(s'|s, y) = 1$ with $s' = (Q_r, \tau_r - 1)$. In this case, the expected reward is $V((Q_r, \tau_r), 0) = \nu - c + \beta V(Q_r, \tau_r - 1)$. On the other hand, if $y = 1$, the user will change to state $s' = (Q_r - Q', \tau_r - 1)$ with probability ω or to state $s' = (Q_r, \tau_r - 1)$ with probability $1 - \omega$. Therefore, we obtain $V((Q_r, \tau_r), 0) = \omega r + \beta \omega V(Q_r - Q', \tau_r - 1) - c(1 - \omega) + \beta(1 - \omega)V(Q_r, \tau_r - 1)$.

Next, we analyze the indexability when $\tau_r > 1$. Let us define

$$\begin{aligned} h(Q_r, \tau_r) &= V((Q_r, \tau_r), 0) - V((Q_r, \tau_r), 1) \\ &= \nu - \omega r + \omega c + \beta \omega(V(Q_r, \tau_r - 1) \\ &\quad - V(Q_r - Q', \tau_r - 1)) \end{aligned} \quad (35)$$

Differentiating $h(Q_r, \tau_r)$ with respect to ν, we obtain $\partial h(Q_r, \tau_r)/\partial \nu = 1 + \beta \omega \partial f(\tau_r - 1)/\partial \nu$. Assuming that $\partial f(\tau_r)/\partial \nu \geq -1/\beta \omega$, we find that $\partial h(Q_r, \tau_r)/\partial \nu \geq 0$. By definition 5, this implies indexability under state (Q_r, τ_r) with $Q_r > 0$ and $\tau_r > 1$. Next, we prove that $\partial f(\tau_r)/\partial \nu \geq -1/\beta \omega$ is true by induction. The expression of $f(\tau_r)$ is

---

Algorithm 1 Online Robustness Control (ORC)

1: **Input:** β, Q_i, $\tau_{max,i}$, c, c'
2: **Initialization:** $Q_r = Q_i$, $\tau_r = \tau_{max,i}$
3: **while** $t < \max\{\tau_{max,i}\}$
4:   **for** $i = 1$: N
5:     SN i probes K^* arms following policy $\hat{\mathcal{P}}_i(t)$ in (25) and sends beliefs ω_k, $k = 1, \ldots, K$ to SO.
6:     SO calculates Whittle index $\nu(s(t)) = \nu(Q_r(t), \tau_r(t))$ as in (37)
7:   **end**
8:   SO activates $(y_k(t) = 1)\gamma B$ connections with highest indexes, and computes (29)
9:   **for** $i = 1$: N
10:     SN i obtains reward (24) based on $y_k(t)$ and updates $Q_r = \sum_k Q_k(t)$, $\tau_r = \tau_{max,i} - t$
11:   **end**
12:   $t = t + 1$
13: **end**

---

calculated as

$$\begin{aligned} f(\tau_r) &= V(Q_r, \tau_r) - V(Q_r - Q', \tau_r) \\ &= \begin{cases} \omega r - (1 - \omega)c \\ \quad + \beta(1 - w)f(\tau - 1), & \text{if } \nu < 0 \\ \omega r - (1 - \omega)c - \nu \\ \quad + \beta(1 - w)f(\tau - 1) & \text{if } 0 \leq \nu < \nu(Q_r, \tau_r) \\ -\omega c + \beta f(\tau - 1), & \text{if } \nu \geq \nu(Q_r, \tau_r) \end{cases} \end{aligned} \quad (36)$$

Since $0 < \beta \leq 1$ and $0 \leq \beta(1 - \omega) < 1$, it is easy to see that $\partial f(\tau_r - 1)/\partial \nu \geq -1/\beta \omega$ for all three cases, which demonstrates the indexability of the connectivity activation problem.

Theorem 1: The connectivity activation problem formulated as a RMAB is indexable.

Proof: See discussion above.

Theorem 2: The closed-form Whittle index $\nu(s)$ of an arm under state $s = (Q_r, \tau_r)$ is

$$\nu(s) = \begin{cases} 0, & \text{if } Q_r = 0 \\ \omega(r + c), & \text{if } Q_r > 0 \\ & \text{and } \tau_r = 1 \\ \omega r \\ \quad + \dfrac{\omega c - \ldots - \beta^{\tau-1}\omega c(1-\omega)^{\tau-1}}{1 - \beta\omega - \ldots - \beta^{\tau-1}\omega(1-\omega)^{\tau-2}}, & \text{if } Q_r > 0 \\ & \text{and } \tau_r > 1 \end{cases} \quad (37)$$

Proof: By definition of the Whittle index, for a given state s, we can obtain the Whittle index by solving (34). From the closed-formed expressions of $V(s, y = 0)$ and $V(s, y = 1)$ previously obtained, and (35)-(36), we have solved (34) and obtained the Whittle index (37).

B. Online Robustness Control Algorithm

We define an Online Robustness Control (ORC) algorithm, as described in Algorithm 1, that combines the online fog

TABLE II

SIMULATION PARAMETERS

Transmission power (P)	23 dBm
Duration of transmission slot (Δt)	10 ms
Transmission bandwidth (W)	10 MHz
Path loss exponent (χ)	4
Message data size (Q)	[0.1 − 1] KB
Maximum latency (τ_{max})	[50 − 100] ms
Noise power spectral density (γ)	−174 dBm/Hz
Message arrival rate (λ_{Pj})	[0 − 1] messages/slot
Computation capacity of the PFN (Γ_j)	[0.7 − 10] GHz
Number of required CPU cycles per message (δ_i)	[0.1 − 1] GHz
Computing energy consumption (ε_i)	$1*10^{-11}$ J/cycle

probing algorithm (25) and the index-based connectivity activation policy (37). At each time slot, the SO calculates the indices of all connections probed by every SN and activates the γB connections with the highest indices. The complexity of calculating all indices is $\mathcal{O}(NK)$ and sorting them has a complexity of $\mathcal{O}(NK\log(NK))$, with $K = (1 + n_c)(1 + n_a)$ where n_c is the number of backup channels and n_a is the number of backup fog nodes. Therefore, the computational complexity of the Whittle index-based activation policy is $\mathcal{O}(NK\log(NK))$.

VII. NUMERICAL RESULTS

In this section, we present numerical results to illustrate the performance of our schemes. The simulations are conducted in Matlab. We simulate a wireless edge/fog network with N SNs, M PFNs, and B channels as described in Section III. The rest of the simulation parameters are summarized in Table II. We simulate three fog probing schemes: online, agnostic and genie. The online fog probing is defined in (25) and builds a belief on the connection availability. The agnostic fog probing is described in Section IV.C and assumes no prior knowledge on the connection availability. Finally, the genie fog probing assumes perfect prediction of connectivity availability and is used for comparison purposes. We simulate the performance of the robustness control framework by using these fog probing schemes and the connectivity activation policy based on Whittle Index and solving the original formulation in (26) with relaxed constraint. First, we evaluate the performance for small-sized networks to compare the results with the original problem formulation (26) that has increasingly exponential complexity with network size and, thus, requires long computational time in large-sized networks. Then, we evaluate the online robustness control algorithm for large-sized networks and show its log-linear scalability. In addition, the results are compared with the least laxity first (LLF) algorithm [33] often used as a benchmark for comparison in scheduling algorithms.

In Fig. 5 we show the robustness of the network ρ^r as in (12) and the average probability of successful transmission (averaged with respect to (8)) for different reinforcement strategies r. We set $N = 20$, $M = 10$ and $B = 5$ (i.e., up to M x $B = 10$ x 5 possible connections per SN). An improvement between 20% to 30% in the robustness was obtained with 1 backup channel and 1 backup PFN compared with no reinforcement, leading to $\rho^r = 0.93$ to 0.995, respectively. Similarly, an improvement of up to 25% was obtained in the average probability of successful transmission under the same
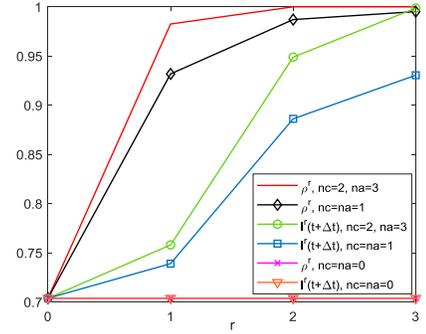
Fig. 5. ρ^r and $l^r(t + \Delta t)$ vs. reinforcement strategy r.
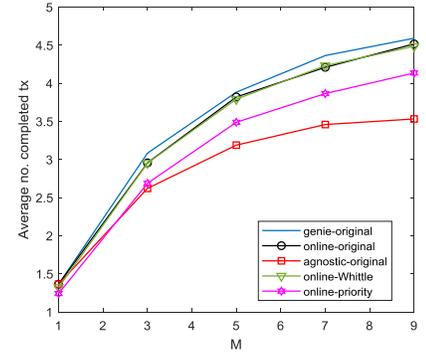
Fig. 6.   Average no. of completed transmissions vs M.

scenario. By increasing the number of backup channels to 2, a robustness $\rho^r = 1$ was obtained and by further increasing the backup PFNs to 3, the average probability of successful transmission equaled to 1.

Next, we conducted Monte Carlo simulations over 10000 realizations of the network to calculate the number of completed transmissions and the reward of the SO. We set $N = 5$, $B = 5, 2$ PNs with $\lambda = 0.5$ and varied M from 1 to 9. Figure 6 shows the average number of completed transmissions (without reinforcement) using the original robustness control formulation (26) with relaxed constraint and the Whittle index for each of the three fog probing schemes, online, agnostic, and genie. In the legend, the first term denotes the fog probing and the second the robustness control algorithm. The genie-aided fog probing corresponds to the case when the SNs can predict with complete certainty the connectivity availability, and thus its performance is the best. The performance of our online fog probing scheme is very close to that of the genie scheme which emphasizes the relevance of our scheme in which the SNs select the connection with the highest belief of availability at the current slot. In the agnostic scheme, the probing starts from a random connection and without prior information on the probability of transmission success. We can see that the online fog probing scheme completed 30% more transmissions than the agnostic scheme. We also compare the SO activation policy solving the original formulation in (26) with the Whittle index policy and a priority policy based on the LLF algorithm [33] in which SNs with more remaining data and fewer slots left had higher priority to transmit. For clarity of presentation, we show the results using the online fog probing. As expected, we observe
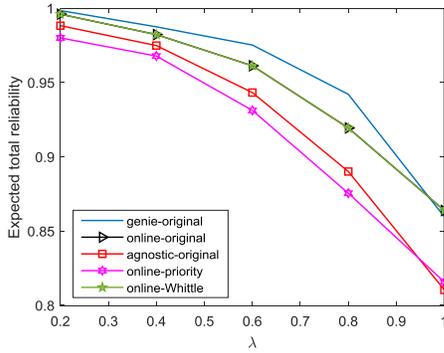
Fig. 7. Expected total long-term reliability vs. λ.
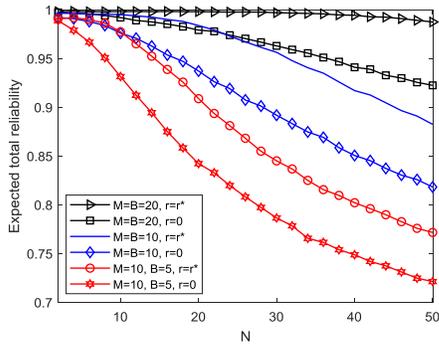
Fig. 8. Expected total long-term reliability vs. N.
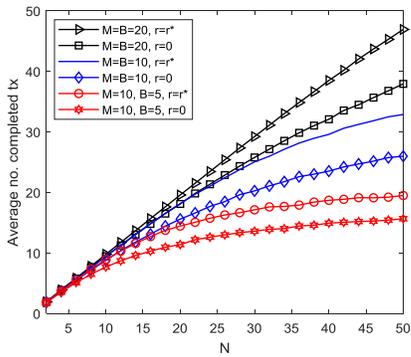
Fig. 9. Average no. of completed transmissions vs. N.

that the Whittle index policy achieved the same performance as solving (26) while the performance under the priority policy is significantly worse. In the latter, the decision is made in each slot without considering the future performance and thus, fewer overall number of transmissions are completed. Similar results were obtained with the other fog probing schemes.

In Fig. 7, the utility of the SO defined as the expected total long-term reliability was obtained for a scenario of $N = M = B = 5$ and 2 PNs for different values of the PNs' traffic arrival rate λ. By increasing λ, the availability of channels and fog nodes decreases. It is worth noting that even for high values of λ (i.e., $\lambda \in [0.6, 0.8]$), the highest deviation of the online fog probing scheme from the genie scheme was about 3%. This demonstrates that our online fog probing algorithm can estimate the belief even when there are many interruptions and so, less past experience available in probing these connections. When $\lambda > 0.8$, there were rarely
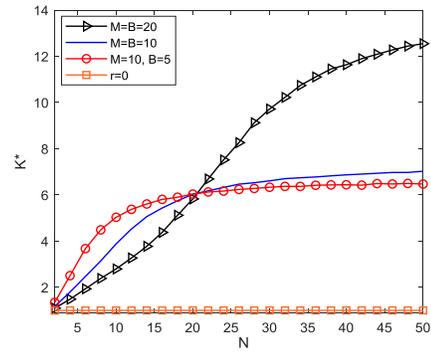
Fig. 10. Optimum number of connections K^* vs. N.

any available connections, so SNs always chose the same connections. As before, the Whittle index activation policy provides the optimum solution to the original problem (26) with relaxed constraint. In the agnostic scheme, under the optimum selection policy, a higher amount of data was scheduled compared with the online-priority scheme. Let us recall that the latter is based on the LLF algorithm [33], which solves the optimization per slot without considering future performance.

Next, we study the performance of the online robustness control algorithm when online fog probing is used together with the Whittle index policy under different reinforcement strategies. In Fig. 8 and Fig. 9, the expected total long-term reliability and the average number of completed transmissions are presented versus N for three scenarios: $M = B = 20$; $M = B = 10$; and $M = 10$, $B = 5$. With the optimum reinforcement strategy r^*, we achieved up to 10% improvement in terms of reliability and completed 15% more transmissions compared to no reinforcement. The optimum number of probed connections K^* to obtain this improvement is shown in Fig. 10. In the first scenario, a total reliability of 0.998 was obtained for $K^* < 6$ probed connections and $N \leq 20$. Since this scenario has the highest number of options for connectivity and backup ($M = B = 20$), the reliability and the number of completed transmissions is significantly higher than in the other two scenarios, especially when $N > 20$. In the second scenario, a total reliability of 0.995 was obtained for 4 probed connections and $N < 10$. In the third scenario, by probing 2.5 connections on average, an expected total reliability of 0.992 was achieved for $N < 7$. By increasing N to 50, the expected total reliability was maximized for 12 and 6 probed connections in the first scenario, and in the second and third scenarios, respectively, achieving a value of 0.987, 0.882 and 0.77. The numbers of completed transmissions in this case, as shown in Fig. 9, were 47, 33 and 20, respectively. Because of interference, fewer SNs could transmit packets simultaneously and, thus, the number of completed transmissions decreased with N. Nevertheless, the improvement with reinforcement strategies was significant in this case as well.

VIII. CONCLUSION AND FUTURE WORK

In this paper we have presented an online robustness control scheme to maximize the total long-term reliability of the connections in a cognitive, dynamic fog/edge-aided wireless network. The scheme consists of two steps. First, a fog probing process is developed to check the availability of the users to act

$$\mathbf{S} = \begin{pmatrix} 0 & (1-\alpha_1)a_{i1}^1 & 0 & \dots & 0 & \alpha_1(1-a_{i1}^1) & (1-\alpha_1)(1-a_{i1}^1) & \dots & 0 & 0 & 0 \\ 0 & 0 & (1-\alpha_2)a_{i2}^1 & \dots & 0 & 0 & \alpha_2(1-a_{i2}^1) & \dots & 0 & 0 & 0 \\ \vdots & & & & & & & & & & \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 \end{pmatrix}$$

(columns labeled) $1 \quad 2 \quad 3 \quad \dots \quad M \quad M+1 \quad M+2 \quad \dots \quad MB-2 \quad MB-1 \quad MB$

$$\mathbf{R} = \begin{pmatrix} \alpha_1 a_{i1}^1 & 0 \\ \alpha_2 a_{i2}^1 & 0 \\ \alpha_3 a_{i3}^1 & 0 \\ \vdots & \vdots \\ \alpha_{M-1} a_{iM-1}^B & 0 \\ \alpha_M a_{iM}^B & 1-\alpha_M a_{iM}^B \end{pmatrix}, \quad \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{0} \in \mathbb{R}^{N_A \times MB}$$

as fog nodes, and the availability of channels and computing resources. An online fog probing algorithm has been presented to maximize the probability that the probed connections are successful (i.e., no transmission outage) during the transmission period. Then, an index-based connectivity activation policy based on RMABs has been proposed. By activating the connections with highest indexes, the total long-term reliability optimization problem is solved with low complexity. Extensive simulations have been conducted to show that our robustness control scheme achieves an expected total reliability very close to the optimum. In addition, it significantly increases the number of completed transmissions compared to an agnostic scheme and a scheme with fixed transmission priorities. Besides, by probing only one additional backup connection, an improvement between 20% to 30% is obtained in terms of network robustness compared with no reinforcement.

In our future work, we will investigate a probable competitive performance ratio by analyzing which of the probed connections the SO activates with a higher probability to reduce the probing cost. Furthermore, we will extend our scenario with heterogeneous IoT devices (SNs) and fog nodes equipped with multiple wireless protocols in different or the same wireless spectrum bands, such as cellular/LoRa (Sub 1GHz), WiFi, Bluetooth, and ZigBee, to collect heterogeneous IoT data. We will study network robustness solutions under heterogeneous traffic patterns and analyze how the traffic changes would affect overall performance.

APPENDIX

State transition probability matrix $\mathbf{S}^{re}$ defined in (13) consist of the matrixes, $\mathbf{S}$ and $\mathbf{R}$ as shown at the top of the page.

REFERENCES

[1] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J.-A. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Commun. Mag.*, vol. 57, no. 8, pp. 84–90, Aug. 2019.

[2] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, May/Jun. 2020.

[3] B. Lorenzo, F. J. Gonzalez-Castano, and Y. Fang, "A novel collaborative cognitive dynamic network architecture," *IEEE Wireless Commun.*, vol. 24, no. 1, pp. 74–81, Feb. 2017.

[4] S. Sarkar, S. Chatterjee, and S. Misra, "Assessment of the suitability of fog computing in the context of Internet of Things," *IEEE Trans. Cloud Comput.*, vol. 6, no. 1, pp. 46–59, Jan./Mar. 2018.

[5] B. Lorenzo, J. Garcia-Rois, X. Li, J. Gonzalez-Castano, and Y. Fang, "A robust dynamic edge network architecture for the Internet of Things," *IEEE Netw.*, vol. 32, no. 1, pp. 8–15, Jan./Feb. 2018.

[6] A. Akbar, M. Ibrar, M. A. Jan, A. K. Bashir, and L. Wang, "SDN-enabled adaptive and reliable communication in IoT-fog environment using machine learning and multiobjective optimization," *IEEE Internet Things J.*, vol. 8, no. 5, pp. 3057–3065, Mar. 2021.

[7] B. Lorenzo, A. S. Shafigh, J. Liu, F. J. Gonzalez-Castano, and Y. Fang, "Data and spectrum trading policies in a trusted cognitive dynamic network architecture," *IEEE/ACM Trans. Netw.*, vol. 26, no. 3, pp. 1502–1516, Jun. 2018.

[8] P. Si, H. Yue, Y. Zhang, and Y. Fang, "Spectrum management for proactive video caching in information-centric cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 8, pp. 2247–2259, Aug. 2016.

[9] Y. Gu, Z. Chang, M. Pan, L. Song, and Z. Han, "Joint radio and computational resource allocation in IoT fog computing," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7475–7484, Aug. 2018.

[10] C. Wang, C. Liang, F. R. Yu, Q. Chen, and L. Tang, "Computation offloading and resource allocation in wireless cellular networks with mobile edge computing," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 4924–4938, Aug. 2017.

[11] X. Chen, "Decentralized computation offloading game for mobile cloud computing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 4, pp. 974–983, Apr. 2015.

[12] J. Li, H. Chen, Y. Chen, Z. Lin, B. Vucetic, and L. Hanzo, "Pricing and resource allocation via game theory for a small-cell video caching system," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 8, pp. 2115–2129, Aug. 2016.

[13] K. H. Abdulkareem *et al.*, "A review of fog computing and machine learning: Concepts, applications, challenges, and open issues," *IEEE Access*, vol. 7, pp. 153123–153140, 2019.

[14] K.-C. Chen, T. Zhang, R. D. Gitlin, and G. Fettweis, "Ultra-low latency mobile networking," *IEEE Netw.*, vol. 33, no. 2, pp. 181–187, Mar./Apr. 2019.

[15] M. Chen, Y. Miao, H. Gharavi, L. Hu, and I. Humar, "Intelligent traffic adaptive resource allocation for edge computing-based 5G networks," *IEEE Trans. Cognit. Commun. Netw.*, vol. 6, no. 2, pp. 499–508, Jun. 2020.

[16] Y. Sun *et al.*, "Adaptive learning-based task offloading for vehicular edge computing systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 3061–3074, Apr. 2019.

[17] J. Yao and N. Ansari, "Fog resource provisioning in reliability-aware IoT networks," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8262–8269, Oct. 2019.

[18] K. Dantu, S. Y. Ko, and L. Ziarek, "RAINA: Reliability and adaptability in Android for fog computing," *IEEE Commun. Mag.*, vol. 55, no. 4, pp. 41–45, Apr. 2017.

[19] J. Yao and N. Ansari, "QoS-aware fog resource provisioning and mobile device power control in IoT networks," *IEEE Trans. Netw. Service Manage.*, vol. 16, no. 1, pp. 167–175, Mar. 2019.

[20] S. Zhao, Y. Yang, Z. Shao, X. Yang, H. Qian, and C.-X. Wang, "FEMOS: Fog-enabled multitier operations scheduling in dynamic wireless networks," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 1169–1183, Apr. 2018.

[21] B. Omoniwa *et al.*, "An optimal relay scheme for outage minimization in fog-based Internet-of-Things (IoT) networks," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3044–3054, Apr. 2019.

[22] Y. Yang, K. Wang, G. Zhang, X. Chen, X. Luo, and M.-T. Zhou, "MEETS: Maximal energy efficient task scheduling in homogeneous fog networks," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 4076–4087, Oct. 2018.

[23] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[24] J. Kwak, O. Choi, S. Chong, and P. Mohapatra, "Processor-network speed scaling for energy–delay tradeoff in smartphone applications," *IEEE/ACM Trans. Netw.*, vol. 24, no. 3, pp. 1647–1660, Jun. 2016.

[25] S.-W. Ko, K. Huang, S.-L. Kim, and H. Chae, "Live prefetching for mobile computation offloading," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3057–3071, May 2017.

[26] B. Lorenzo, I. Kovacevic, A. Peleteiro, F.-J. Gonzalez-Castano, and J. C. Burguillo, "Joint resource bidding and tipping strategies in multi-hop cognitive networks," *IEEE Trans. Cognit. Commun. Netw.*, vol. 2, no. 3, pp. 301–315, Sep. 2016.

[27] S. Glisic, B. Vucetic, *Spread Spectrum CDMA Systems for Wireless Communications*. Norwood, MA, USA: Artech House, 1997.

[28] K. Liu and Q. Zhao, "Indexability of restless bandit problems and optimality of whittle index for dynamic multichannel access," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5547–5567, Nov. 2010.

[29] J. Xu and C. Guo, "Scheduling periodic real-time traffic in lossy wireless networks as restless multi-armed bandit," *IEEE Wireless Commun. Lett.*, vol. 8, no. 4, pp. 1129–1132, Aug. 2019.

[30] A. Brzezinski, G. Zussman, and E. Modiano, "Distributed through-put maximization in wireless mesh networks via pre-partitioning," *IEEE/ACM Trans. Netw.*, vol. 16, no. 6, pp. 1406–1419, Dec. 2008.

[31] S. Hosseinalipour, C. G. Brinton, V. Aggarwal, H. Dai, and M. Chiang, "From federated to fog learning: Distributed machine learning over heterogeneous wireless networks," *IEEE Commun. Mag.*, vol. 58, no. 12, pp. 41–47, Dec. 2020.

[32] J. Xu and C. Guo, "Scheduling stochastic real-time D2D communications," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 6022–6036, Jun. 2019.

[33] Z. Yu, Y. Xu, and L. Tong, "Deadline scheduling as restless bandits," *IEEE Trans. Autom. Control*, vol. 63, no. 8, pp. 2343–2358, Aug. 2018.

[34] S. Liao, J. Wu, S. Mumtaz, J. Li, R. Morello, and M. Guizani, "Cognitive balance for fog computing resource in Internet of Things: An edge learning approach," *IEEE Trans. Mobile Comput.*, early access, Sep. 24, 2020, doi: 10.1109/TMC.2020.3026580.

[35] I. A. M. Balapuwaduge, F. Y. Li, A. Rajanna, and M. Kaveh, "Channel occupancy-based dynamic spectrum leasing in multichannel CRNs: Strategies and performance evaluation," *IEEE Trans. Commun.*, vol. 64, no. 3, pp. 1313–1328, Mar. 2016.

[36] Q. Liang, X. Wang, X. Tian, F. Wu, and Q. Zhang, "Two-dimensional route switching in cognitive radio networks: A game-theoretical framework," *IEEE/ACM Trans. Netw.*, vol. 23, no. 4, pp. 1053–1066, Aug. 2015.

[37] P.-K. Tseng and W.-H. Chung, "Local rerouting and channel recovery for robust multi-hop cognitive radio networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2013, pp. 2895–2899.

[38] A. Maatouk, S. Kriouile, M. Assad, and A. Ephremides, "On the optimality of the Whittle's index policy for minimizing the age of information," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 1263–1277, Feb. 2021.

Francisco Javier González-Castaño is currently a Full Professor with the Telematics Engineering Department, University of Vigo, Spain, where he leads the Information Technology Group. He has authored over 100 articles in international journals, in the fields of telecommunications and computer science, and has participated in several relevant national and international projects. He holds three U.S. patents.

Linke Guo (Senior Member, IEEE) received the B.E. degree in electronic information science and technology from the Beijing University of Posts and Telecommunications in 2008, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Florida, FL, USA, in 2011 and 2014, respectively. From August 2014 to August 2019, he was an Assistant Professor with the Department of Electrical and Computer Engineering, Binghamton University. Since August 2019, he has been an Assistant Professor with the Department of Electrical and Computer Engineering, Clemson University. His research interests include wireless networks, the IoT, security, and privacy. He is currently serving as an Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He also serves as the Poster/Demo Chair for IEEE INFOCOM (2020–2021).

Felipe Gil-Castiñeira is currently an Associate Professor with the Department of Telematics Engineering, University of Vigo. His research interests include wireless communication and core network technologies, multimedia communications, embedded systems, ubiquitous computing, and the Internet of Things. He has published over 60 articles in international journals and conference proceedings. He has led several national and international research and development projects. He holds two patents in mobile communications.

Yuguang Fang (Fellow, IEEE) received the M.S. degree from Qufu Normal University, China, in 1987, the Ph.D. degree from Case Western Reserve University, Cleveland, OH, USA, in 1994, and the Ph.D. degree from Boston University, Boston, MA, USA, in 1997. He joined the Department of Electrical and Computer Engineering, University of Florida, FL, USA, in 2000, where he has been a Full Professor since 2005 and a Distinguished Professor since 2019. He received the US NSF CAREER Award in 2001, the US ONR Young Investigator Award in 2002, the 2015 IEEE Communications Society CISTC Technical Recognition Award, the 2018 IEEE Vehicular Technology Outstanding Service Award, and the 2019 IEEE Communications Society AHSN Technical Achievement Award. He has been actively participating in conference organizations such as serving as the Technical Program Co-Chair for IEEE INFOCOM'2014 and the Technical Program Vice-Chair for IEEE INFOCOM'2005. He has been serving on several editorial boards for journals, including Proceedings of the IEEE since 2018, ACM Computing Surveys since 2017, *ACM Transactions on Cyber-Physical Systems* since 2020, IEEE TRANSACTIONS ON MOBILE COMPUTING since 2019. He was serving on several editorial boards for journals, including IEEE TRANSACTIONS ON MOBILE COMPUTING from 2003 to 2008 and from 2011 to 2016, IEEE TRANSACTIONS ON COMMUNICATIONS from 2000 to 2011, and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2002 to 2009.

Beatriz Lorenzo (Senior Member, IEEE) received the M.Sc. degree in telecommunication engineering from the University of Vigo, Spain, in 2008, and the Ph.D. degree from the University of Oulu, Finland, in 2012. She is currently an Assistant Professor with the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst. Her research interests include AI for wireless networks, network architectures and protocol design, mobile computing, optimization, and network economics. From 2016 to 2017, she received the Fulbright Visiting Scholar Fellowship with the University of Florida. She is an Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. She served as Co-Chair for WiMob Conference in 2019.