Implementation of an Age Prioritized Wi-Fi Network with Multi-Armed Bandit Polling

Alperen Balci, Batu Saatci, Emir Atak, and Elif Uysal
Communication Networks Research Group (CNG),
Electrical and Electronics Eng. Dept, METU, Ankara, Turkey

***Abstract*—This paper presents a method for prioritizing transmissions considering the Age of Information (AoI) metric in communication systems that are compatible with the IEEE 802.11 standard. The proposed method includes Last-In-First-Out (LIFO) at users and a multiuser scheduling algorithm that uses a Multi-Armed Bandit (MAB) model at the access point. The system was implemented on ESP32 microcontrollers and Raspberry Pi devices to evaluate performance, making our study one of the few real-life implementations of age-aware WiFi. The performance of the proposed method is compared with that of the WiFresh method, which uses a Maximum Weight matching scheduler. The simulation and implementation results indicate that the proposed method improves the Peak and mean AoI in the system compared to Wi-Fi. The implementation demonstrates the applicability of the method in real-world scenarios.**

***Index Terms*—Wi-Fi, Age of Information, IEEE802.11, Multi-Armed Bandit, Information Freshness, Semantic Communication**

I. Introduction

Wireless local area networks adopting IEEE802.11 family of standards [1], sometimes called Wi-Fi, have revolutionized wireless connectivity, enabling seamless connectivity and data exchange among devices for personal and device communication in indoor and outdoor environments. Today, especially in urban areas, numerous Wi-Fi networks are within range at any given location. With the proliferation of Internet of Things (IoT) applications in recent years, wireless access has been increasingly utilized for machine-type communication, file transfer, and multimedia. This type of communication involves the intermittent transmission of small data packets. Often, particularly in applications involving status updates, transmitted data remain useful only for a limited period, making data freshness a crucial factor. Moreover, the density of such transmitters can be 10-100 times higher than that of traditional user terminals. This situation highlights a critical scalability issue, where data packets experience prolonged delays in buffers due to limited access to the medium shared among Wi-Fi users, leading to outdated data, and causing unnecessary congestion and inefficiency in the network [2].

Such issues raise the need to define a new performance indicator called the Age of Information (AoI). AoI of a flow can be defined as the time elapsed since the newest received packet available at the destination was generated [3] at the source. Scheduling policies that minimize AoI have been proposed to improve the freshness of data packets in communication systems [4] and a Wi-Fi design utilizing this principle has been suggested as a solution [5]. In this study, we propose and implement a new method to prioritize transmissions by considering the AoI metric in communication systems compatible with the IEEE 802.11 standard.

This work was partly supported by TUBITAK grant 22AG019.

With growing interest in data freshness within communication systems, research on Age of Information (AoI) has expanded significantly. Practical considerations of AoI are discussed in [6], while [7] presents an experimental evaluation of an age-aware protocol for IoT applications. Additionally, [8] introduces an Age-Aware Application Layer Forward Error Correction Flow Control mechanism. Age-aware maximum weight matching was considered in [5], [9], age optimization was employed in association with specific channel access methods in [4], and a study aimed at channel learning in [10]. WiFresh has shown superior age performance at the application layer compared to standard Wi-Fi communication [5]. This distinction becomes more evident as the network traffic grows, raising some questions regarding the scalability of such wifi networks. In this study, a LIFO queue structure is used at users instead of the FIFO queue structure in standard Wi-Fi. Additionally, user data is controlled using a Multi-Access scheduling mechanism with the Maximum Weight (MW) principle.

It has been observed [11] that replacing the usual FIFO service discipline with a LIFO link policy delivers most of the potential age improvement possible. When freshness is the sole performance indicator, it is not advantageous to deliver older packets. This means that the optimal decision is to keep only the freshest data in the queue and replace it with the freshest data upon arrival. However, dropping old packets can be a disadvantage in systems where historical data analysis is desired. Detailed studies have been conducted in [11] and [12].

The Multi-Access scheduling mechanism can utilize the Multi-Armed Bandit formulation. This machine learning branch aims to learn to select the most rewarding choice among many options [13]. Each option has a previously unknown reward, and the goal is to maximize the reward. In a study on an LP-WAN structure with an AoI-prioritized system, the MAB methodology was used in delay-tolerant systems, yielding better results in terms of AoI [10] compared to other algorithms. In systems where the trend of the reward function in users

979-8-3315-2096-0/25/$31.00 ©2025 IEEE

is known, additional improvements to the MAB solution have shown positive results [14].

In this study, the structure in [5] is further developed and its experimental performance is demonstrated. A new MAB solution proposed and implemented on ESP32 microcontrollers and Raspberry Pi devices is proposed. The performance of the proposed method is compared with that of standard Wi-Fi and the WiFresh method.

II. SYSTEM DESIGN

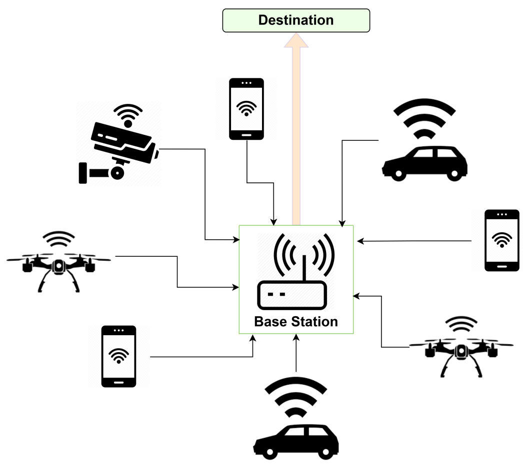

Fig. 1: System Model

In this paper, a Wi-Fi network structure is discussed. In this wireless network structure, various units (users) aim to send information to a central access point (AP), and the AP forwards this information to the managing unit. The basic structure of the WiFresh network is shown in Figure 1.

Users update the information they want to send after a defined time interval. According to the AP scheduling policy, AP sends a poll packet (PP) to the user at each suitable time slot to prompt them to send their packet. The user who receives the PP sends their queued packet to the AP. Users also add the generation time of the packet to their packets. The AP records the AoI value for each user. The AoI for each user i, denoted by $\Delta_i(t)$, is calculated as shown in (1).

$$\Delta_i(t+1) = \begin{cases} 1 & \text{, if packet is received from i at } t \\ \Delta_i(t)+1 & \text{, otherwise} \end{cases} \tag{1}$$

The objective of the optimization is to find a scheduling policy π that minimizes the average AoI at the destination and $J_A(\pi)$ is the average AoI incurred while using policy π. This optimization problem can be formulated as follows:

$$\min_{\pi\in\Pi} \mathbb{E}\big[J_A(\pi)\big], \quad \text{where} \quad J_A(\pi) = \frac{1}{TN}\sum_{t=1}^{T}\sum_{i=1}^{N}\Delta_i(t). \tag{2}$$

In [15], a lower bound for this optimization problem is derived. Here, p_i represents the probability of successful transmission (channel reliability) for node i, T is the total time and N is the number of users. When all users have the same weight, the lower bound is expressed as:

$$\lim_{T\to\infty} J_A(\pi) \geq \frac{1}{2N}\left(\sum_{i=1}^{N}\sqrt{\frac{1}{p_i}}\right)^2 + \frac{1}{2} \tag{3}$$

The lower bound is achieved when the AP selects the user with the highest AoI at each time slot. The following sections will discuss policies to convert this optimization problem into a practical implementation.

III. AGE OF INFORMATION PRIORITIZED WI-FI

In standard Wi-Fi, the FIFO queue structure (First-In-First-Out) is used. In this structure, the freshest packet is added to the back of the queue and must wait for all packets at the front to be transmitted. When data rates are insufficient, the freshest packet experiences high queueing delays, and data becomes stale. So, LIFO (Last-In-First-Out) queues are used in our suggested AoI-prioritized system to prevent this situation.

A random access mechanism in standard Wi-Fi coordinates users sharing one common channel. This mechanism resolves packet collisions in the channel by letting users with colliding packets wait for a random backoff. However, as the number of users increases, the probability of packet collisions also increases, which results in longer back-off times and longer transmission delays.

The multi-access mechanism prevents packet collision by picking which user can send a packet to the channel. Such a mechanism needs a selection policy to be defined, which chooses users according to some priorities determined beforehand. In the following subsections, we will analyze the Maximum Weight policy described in [5] and the proposed MAB structure as selection policies.

A. Maximum Weight

The AP uses the Age-Based Maximum Weight (MW) principle to decide which user to select. MW calculates a priority value for each user based on the AoI and the user's channel reliability and picks the user with the highest priority. At time t, the priority value of user i is calculated as shown in (4).

$$I(i,t) = p_i(\Delta_i(t) - H_i(t))^2 \tag{4}$$

Where p_i is the channel reliability of user i and takes a value in the range $(0, 1]$, $\Delta_i(t)$ is the AoI of user i at time t as seen by the AP shown in (5).

$$\Delta_i(t) = t - \tau_i(t) \tag{5}$$

Where $\tau_i(t)$ is the last time the freshest data from user i was obtained before time t. The parameter $H_i(t)$ in (4) is the age of the data that will be obtained if the user i is selected as seen in (6).

$$H_i(t) = t - \tau_i^{HoL}(t) \tag{6}$$

$\tau_i^{HoL}(t)$ represents the creation time of the freshest data that user i has at time t. The AP selects the user with the highest $I(i,t)$ values at each decision-making moment. However, to apply the principle of MW, the values of p_i, $H_i(t)$, and $\Delta_i(t)$ must be estimated at the AP. This is because the AP does not know the instantaneous AoI and channel reliability values of the users at the selection time t. The estimation of $\hat{p}_i(t)$ is done as follows.

$$\hat{p}_i(t) = \frac{D_i(t)+1}{P_i(t)+1} \tag{7}$$

In (7), $P_i(t)$ represents the number of times the user i has been selected. In contrast, $D_i(t)$ represents the number of successful transmissions to the AP by the user i during these selections. The $+1$ values in the (7) ensure that the channel reliability of users who have never transmitted before is estimated as 1. This increases the likelihood of allowing newly added senders to transmit after being added to the system. For this estimation, the values of $D_i(t)$ and $P_i(t)$ are recorded during implementation.

Instead of using the transmission and acquisition of the time information of all the packets received up to that point, only the information of the most recently received packet is used to estimate the value of $H_i(t)$, $\hat{H}_i(t)$. In this estimation, the AP equates the user's age to the age of the most recent packet received from the user. This age is not changed unless a new packet is received. This ensures that the probability of a user who has recently made a transmission making another transmission is initially very low but increases over time.

B. Polling Mechanism with MAB

It has been observed that the AP's estimations of the Age of Information and channel reliability for users and the selections made based on these estimations can be formulated as a Multi-Armed Bandit problem.

Upon the delivery of the packet to the AP, the reduction in the Age of Information for the corresponding user at the AP is calculated by subtracting the value of $H_i(t)$ from the current value of $\Delta_i(t)$, considering the packet delay.

The reward function is the reduction in the Age of Information of the data sent by the user at that moment. The reward function is given by (8).

$$R_i(t+1) = \begin{cases} \Delta_i(t) - H_i(t) & \text{, if transmission occurs} \\ 0 & \text{, otherwise} \end{cases} \tag{8}$$

We aim to keep the Age of Information in the system as low as possible by maximizing the total reward. When the selection policy is expressed as $\pi \in \Pi$ and considering that there are N different senders over a total time T, the situation to be maximized can be expressed as follows.

TABLE I: SIMULATION RESULTS

N:	4	8	12	16	20	
MAoI (s)	0.000917	0.0021	0.0759	2.0209	9.6107	WiFi
	0.000851	0.0019	0.0029	0.0038	0.0048	WiFresh
	0.000737	0.0016	0.0025	0.0034	0.0043	UCB-PI
PAoI (s)	0.0032	0.0051	0.1534	3.4702	18.4223	WiFi
	0.0038	0.0054	0.0064	0.0076	0.0084	WiFresh
	0.0033	0.0043	0.0061	0.0063	0.0077	UCB-PI

$$\max_{\pi \in \Pi} \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N} E[R_i(t)] \tag{9}$$

Since the AP does not know the reward it may earn from polling any user at any given time, it has to estimate it. For this purpose, the Upper Confidence Bound (UCB) has been used [13]. Because the reward function in users has a known behavior, the A-UCB structure proposed in [14] has been modified for AoI-sensitive systems, and the reward function has been added as a multiplier to the standard UCB structure. We will refer to this structure, given by (10), as UCB-PI (UCB with Potential Information).

$$UCB_{i,t} := \left(\frac{\sum_{k=1}^{t} R_i(k)}{n_{i,t}} + c \times \sqrt{\frac{2\ln(t)}{n_{i,t}}}\right) \times (\Delta_i(t) - H_i(t)) \tag{10}$$

In (10), the value $n_{i,t}$ denotes the total number of packets sent by the sender i up to time t. The parameter $\sqrt{\frac{2\ln(t)}{n_{i,t}}}$ is used to prevent the same user from transmitting continuously. Subsequently, at each time t, the users with the highest UCB values can transmit packets to the AP.

IV. SIMULATIONS AND IMPLEMENTATIONS

A. Simulation

MATLAB simulations have been conducted to compare the average Age of Information (MAoI) at the AP and the maximum Age of Information (PAoI) in the system among standard Wi-Fi, WiFresh, and UCB-PI.

The simulations were conducted with the following parameters: the system was set to N users, with PPs sent to 1 user at each decision time. The probability of successful transmission for each user was randomly selected from the range $(0.5, 1]$ and was not changed during the simulation. The results are presented in Table (I).

It has been observed that MAB reduces the average and maximum AoI by about a factor of two thousand, compared to standard Wi-Fi. Hence, it brings a three-order improvement. Moreover, it also achieves an approximately 10% improvement over WiFresh. Therefore, the MAB structure is more effective in reducing the mean and peak AoI in the system compared to WiFresh and standard Wi-Fi.

B. Implementation

The proposed system was implemented by making modifications at the application layer due to cost considerations and ease of integration. This allowed the system to be added

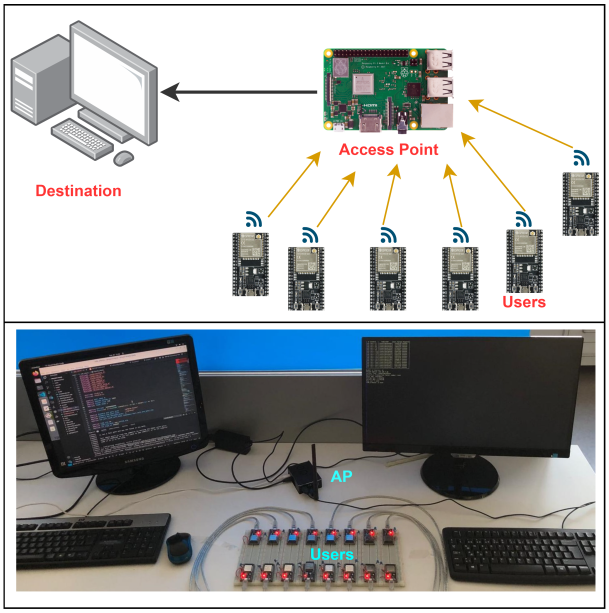

Fig. 2: Implementation Setup

to the standard UDP, IP, and Wi-Fi structures. ESP32s were used as users, while a Raspberry Pi-4 with higher processing capacity was used as the AP as shown in figure (2). A personal computer was used as the management unit. The system was configured to operate on the IEEE802.11b standard with a central frequency of 2.4 GHz. The packet sizes were set to 150 bytes. The system was set to operate with 4, 8, 12, and 16 users. The users were placed in the same room as the AP, and the AP was connected to the managing unit via Ethernet.

For WiFresh and UCB-PI, the user's channel reliability, potential AoI reduction estimates, and historical reward information were stored in the Raspberry Pi-4's memory. A timeout period was started when a selection packet was sent to the chosen user. The AP updated its estimates based on whether the user received a response within this period. When users send their packets, they stamp them with the generation time. The AP calculated the AoI of the users based on this information. Each experiment was conducted until 10000 packet transmissions were completed.

The most challenging part of the implementation was the synchronization of the users between themselves and the AP. Time synchronization between users and the AP was achieved using the Network Time Protocol (NTP). The period determined for NTP synchronization was set to 15 seconds. Moreover, the ESP32 internal clock was volatile and the clock drift was very high. Therefore, we used an external very low-jitter oscillator instead of ESP32's internal clock to achieve better synchronization and reduce the clock drift.

The algorithm implemented in the AP and its pseudocode is provided in Algorithm 1 and the algorithm implemented in users and its pseudocode is provided in Algorithm 2.

TABLE II: Comparison of Mean AoI and Peak AoI for different N values

N	WiFresh			MAB - UCB-PI		
	MAoI	PAoI	Throughput	MAoI	PAoI	Throughput
4	0.008	0.046	0.621	0.009	0.049	0.638
8	0.012	0.061	0.595	0.012	0.064	0.596
12	0.016	0.074	0.594	0.015	0.077	0.615
16	0.021	0.082	0.555	0.020	0.079	0.600

TABLE III: Mean AoI, Peak AoI, and Throughput Values for different numbers of users in standard WiFi

N	MAoI	PAoI	Throughput
4	0.30	0.62	1.38
8	0.97	2.81	1.09
12	1.50	3.54	0.95
16	2.26	4.65	0.74

The implementation results for 16 users where the 1KHz data generation rate and the channel reliability rate 1 are presented in Tables II and III. Throughputs are given in Mbps. The results show that the UCB-PI structure reduces the system's maximum AoI and mean AoI more effectively than the wifresh and standard Wi-Fi when user number is 16. When the number of users increases, the UCB-PI structure provides better results regarding the maximum AoI and mean AoI. In addition, the UCB-PI structure has better throughput than WiFresh.

Algorithm 1 Multi-Armed Bandit Algorithm for Access Point

Initialize: N arms (users), T time steps, $n_i = 0$ for all $i \in \{1, \dots, N\}$, $R_i = 0$ for all $i \in \{1, \dots, N\}$
while true **do**
    for each arm $i \in \{1, \dots, N\}$ **do**
        if $n_i > 0$ **then**
            Calculate $UCB_{i,t}$ use Eq. (10)
        else
            Set $UCB_{i,t} = \infty$ ▷ Try each arm at least once
        end if
    end for
    Select arm $i^* = \arg\max_i UCB_{i,t}$
    Transmit poll packet to node i^*
    if timeOut = true **then**
        Update $n_{i^*} = n_{i^*} + 1$
    else if packet received from i^* **then**
        Receive reward $r_{i^*}(t)$
        Update $n_{i^*} = n_{i^*} + 1$
        Update $R_{i^*} = R_{i^*} + r_{i^*}(t)$
    end if
end while

The comparison between Wifresh and UCB-PI for 16 users, channel reliability 0.75, and period 1,10,100,1000 Hz for every 4 users sequentially is presented in Table IV. The results show that the UCB-PI structure is more effective in reducing the

Algorithm 2 Algorithm for Users

```
Initialize: Synchronize time with the access point, set data period B, set data generation probability k
Variables: packetStack ← ∅, currentTime ← 0, B ← data period, k ← data generation probability
Procedure: On entering the system
Synchronize time with the access point
Add node to the system
while node is active do
    currentTime ← getCurrentTime()
    if currentTime mod B == 0 then
        if random() < k then
            Generate a new packet with timestamp currentTime
            Push the new packet onto packetStack
        end if
    end if
    if PP received from access point then
        if packetStack is not empty then
            latestPacket ← pop from packetStack
            Send latestPacket to access point
        end if
    end if
end while
```

maximum AoI in the system compared to the wifresh when users have different data generation rates.

TABLE IV: Results for N=16, p=0.75, Period 1,10,100,1000 Hz for each 4 users sequentially

N	WiFresh			MAB - UCB-PI		
	MAoI	PAoI	Throughput	MAoI	PAoI	Throughput
16	1.1	7.8032	0.1054	1.103	7.156	0.1038

The implementation results did not align fully with the simulations. The discrepancies are believed to be due to the processing time in the AP and the users. The reason is mainly due to the lower processing power of the ESP32s. Additionally, a higher maximum AoI observed in standard Wi-Fi is due to users sometimes being unable to transmit data because of packet collisions. This has limited the maximum speed the system can achieve.

V. Conclusion

This study on prioritizing the Age of Information (AoI) in Wi-Fi networks introduces a new method within the framework of existing IEEE802.11 standards, aiming to improve communication freshness using the AoI concept. The study uses a hardware setup to present a new Multi-Armed Bandit (MAB) algorithm named UCB-PI to prioritize AoI in Wi-Fi networks. The UCB-PI algorithm is compared with the WiFresh algorithm, which is based on the Maximum Weight principle and also compares the performance of the UCB-PI and WiFresh algorithms with standard Wi-Fi in real implementations.

Simulation and implementation studies show that the UCB-PI and WiFresh algorithms improve the average and maximum AoI compared to standard Wi-Fi. The UCB-PI algorithm reduces the system's maximum AoI more effectively than WiFresh. Additionally, the UCB-PI algorithm has better throughput compared to WiFresh. The results show that the UCB-PI algorithm reduces the system's maximum AoI more effectively than WiFresh when users have different data generation rates.

Improving the real-life performance of the UCB-PI algorithm by addressing practical implementation issues is left to future work. Other directions for further work include investigation of the impact of different channel conditions and different packet sizes on the performance of the MAB algorithm.

Acknowledgement

References

[1] M. S. Gast, *802.11 Wireless Networks: The Definitive Guide*. O'Reilly Media, 2017.

[2] E. Uysal, O. Kaya, A. Ephremides, J. Gross, M. Codreanu, P. Popovski, M. Assaad, G. Liva, A. Munari, T. Soleymani, B. Soret, and K. Johansson, "Semantic communications in networked systems," *arXiv:2103.05391*, 2021.

[3] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *Proc. IEEE Commun. Soc. Conf. Sensor, Mesh, and Ad Hoc Commun. and Netw. (SECON)*, 2011, pp. 350–358.

[4] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Scheduling policies for minimizing age of information in broadcast wireless networks," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2637–2650, Dec. 2018.

[5] I. Kadota, M. S. Rahman, and E. Modiano, "WiFresh: Age-of-Information from Theory to Implementation," in *Proc. ACM MobiCom*, 2020.

[6] E. Uysal, O. Kaya, S. Baghaee, and H. B. Beytur, "Age of Information in Practice," *CoRR*, vol. abs/2106.02491, 2021. [Online]. Available: https://arxiv.org/abs/2106.02491.

[7] U. Guloglu, S. Baghaee, and E. Uysal, "Evaluation of Age Control Protocol (ACP) and ACP+ on ESP32," in *Proc. 2021 17th Int. Symp. Wireless Commun. Syst. (ISWCS)*, 2021, pp. 1–6. doi: 10.1109/ISWCS49558.2021.9562140.

[8] S. Baghaee and E. Uysal, "A^3L-FEC: Age-Aware Application Layer Forward Error Correction Flow Control," *arXiv preprint arXiv:2410.05852*, 2024. [Online]. Available: https://arxiv.org/abs/2410.05852.

[9] J. Sun, Z. Jiang, B. Krishnamachari, S. Zhou, and Z. Niu, "Closed-form Whittle's index-enabled random access for a timely status update," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1824–1837, Mar. 2019.

[10] O. Gencay, A. Balci, E. T. Ceran, and E. Uysal, "Ensuring a semantically effective IoT network through blockchain with delayed feedback," in *Proc. Int. Conf. Comput., Netw. and Commun. (ICNC)*, 2024.

[11] A. M. Bedewy, Y. Sun, and N. B. Shroff, "The age of information in multihop networks," *IEEE/ACM Trans. Netw.*, vol. 27, no. 2, pp. 1248–1261, Apr. 2019.

[12] S. Kaul, R. D. Yates, and M. Gruteser, "Status updates through queues," in *Proc. IEEE CISS*, 2012, pp. 1–6.

[13] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire, "The nonstochastic multiarmed bandit problem," *SIAM J. Comput.*, vol. 32, no. 1, pp. 48–77, Jan. 2002.

[14] D. Bouneffouf and R. Feraud, "Multi-armed bandit problem with known trend," *arXiv:1508.07091*, 2017.

[15] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Scheduling policies for minimizing age of information in broadcast wireless networks," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2637–2650, Dec. 2018.