

# Traffic Prediction and Fast Uplink for Hidden Markov IoT Models

Elslam Eldeeb, Mohammad Shehab, Anders E. Kalør, Petar Popovski, and Hirley Alves

**Abstract**—In this work, we present a novel traffic prediction and fast uplink framework for IoT networks controlled by binary Markovian events. First, we apply the forward algorithm with hidden Markov models (HMM) in order to schedule the available resources to the devices with maximum likelihood activation probabilities via fast uplink grant. In addition, we evaluate the regret metric as the number of wasted transmission slots to evaluate the performance of the prediction. Next, we formulate a fairness optimization problem to minimize the age of information while keeping the regret as minimum as possible. Finally, we propose an iterative algorithm to estimate the model hyperparameters (activation probabilities) in a real-time application and apply an online-learning version of the proposed traffic prediction scheme. Simulation results show that the proposed algorithms outperform baseline models such as time division multiple access (TDMA) and grant-free (GF) random-access in terms of regret, the efficiency of system usage, and age of information.

**Index Terms**—Age of information, fast uplink, hidden Markov model, internet-of-things, online learning, resource allocation.

## I. INTRODUCTION

Recent advances in internet of things (IoT) has led to the deployment of a large number of machine-type communication (MTC) devices to collect real-time information. The number of such IoT-MTC devices is rapidly growing to realize different use cases such as environment monitoring, remote surgery, and autonomous vehicles [1]. In 5G, MTC service modes are massive MTC (mMTC) and ultra-reliable low latency communication (URLLC) [2]. The quality-of-service (QoS) demands vary among the service modes. In addition, many use cases have recently had more strict demands, which need extremely low end-to-end latency in a massive deployment of IoT devices to collect real-time information [2].

The behavior of the traffic of MTC devices (MTDs) differs from that of the traditional human-type communication devices (HTDs) [3]. The HTDs traffic tends to be heterogeneous, whereas the traffic of MTDs is homogeneous and highly correlated. To elucidate traffic correlation in MTC, we consider the following road safety example as in Fig. 1: let event 1

The work by E. Eldeeb, M. Shehab, and H. Alves has been partially supported by Academy of Finland 6Genesis Flagship (Grant no. 318927), and FIREMAN (Grant no. 326301), and the European Commission through the Horizon Europe project Hexa-X (Grant Agreement no. 101015956). The work by A. E. Kalør and P. Popovski has been supported by the Danish Council for Independent Research (Grant Nr. 8022-00284B SEMIOTIC) and by the Villum Investigator Grant "WATER" from the Velux Foundation, Denmark.

E. Eldeeb, M. Shehab, and H. Alves are with Centre for Wireless Communications (CWC), University of Oulu, Finland. Email: first-name.lastname@oulu.fi.

A. E. Kalør and P. Popovski are with the Department of Electronic Systems, Aalborg University, Aalborg, Denmark e-mails: aek, petarp@es.aau.dk.

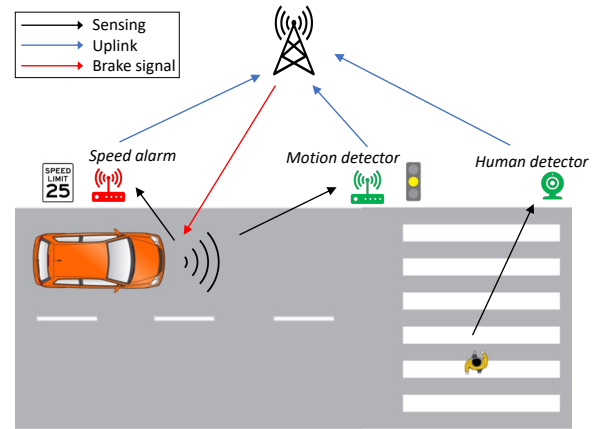


Fig. 1: Traffic correlation scenario: A speed alarm would be active only if the motion detector is active but not vice versa. The human detector signal would only be important if the other two sensors are active.

and event 2 correspond to a vehicle moving down the street at normal speed, and a vehicle breaking the speed limit, respectively. Meanwhile, sensor 1 and sensor 2 are motion detectors, necessary to control the traffic lights, and speed limit alarm, respectively. In this scenario, event 1 will be detected by sensor 1 only. However, both sensors may likely detect event 2. Hence, we infer that sensor 2 will not likely be active except if sensor 1 is active. Moreover, if sensor 2 is active, sensor 1 will most probably be active but not vice versa. In such a scenario, it is essential to estimate the possible sensor activation pattern and allocate resources at low latency. If a human is crossing the street, a human detector or a road safety alarm could then transmit a signal to the BS. The BS in turn sends a compulsory brake signal to a high-speed vehicle to enforce it to slow down the speed. This all should occur within a window of a few milliseconds to avoid an accident. The importance of an uplink signal from the human detector in this scenario is also dependant on whether the speed alarm is active or not.

Another example to illustrate traffic correlation, let Markovian event 1 and Markovian event 2 correspond to the existence of fire or no fire, and someone who smokes a cigarette or no smoke, respectively. Meanwhile, sensor 1 and sensor 2 are heat and smoke detectors, respectively. In case of fire, both sensors will detect the event. However, in case of smoking a cigarette, event 2 will only be detected by sensor 2. Hence, we infer that if sensor 1 is active, sensor 2 will be active with high probability but not vice versa.

One important metric to measure the freshness of received data from an IoT device is the age-of-information (AoI). AoI was first introduced in [4]. It defines the freshness of information (time elapsed since data at a source has been collected and transmitted to a destination). Therefore, minimizing the AoI in IoT networks has become essential when designing scheduling algorithms [5].

A key element in communication systems is the design of access protocols, which allow the devices to transmit their data in an organized manner. In what follows, we discuss the shortcomings of existing massive access protocols. In conventional LTE systems, the devices communicate with the base station (BS) using the random access (RA) procedures [6], e.g., each device goes through a 4-handshake procedure initiated by the transmission of a random preamble followed by a random-access response from the BS side. Afterward, the device requests a connection and the BS responds with a contention resolution message. However, this procedure suffers from high signaling overhead and end-to-end latency, which fails to serve strict low latency demands and results in a relatively high AoI. Furthermore, due to the limited number of preambles, it is susceptible to a high number of collisions in situations where a large number of devices sporadically try to access the network at the same time, such as in an alarm scenario [7].

Meanwhile, alternative solutions have been proposed to solve the problems of collisions and signaling overhead in IoT networks, from legacy time division multiple access (TDMA) to grant-free (GF) schemes, and access class barring (ACB). In TDMA schemes, the resources are distributed equally among the devices without considering any scheduling algorithms. Although TDMA is straightforward and efficient in periodic transmission scenarios, it does not perform well when the traffic is sporadic and event-driven. Therefore, GF access has been proposed as an efficient procedure to reduce the signaling overhead by skipping the preamble request and reply that constitute the first 2-steps of the 4-handshake procedure [8]. Although GF solutions reduce the signaling overhead to half of the grant-based RA, it fails when the number of potentially active devices exceeds the available resources. In addition, it suffers from a large number of collisions, which cause high AoI experienced by the devices. Among the alternative solutions, promising results have been obtained for ACB [9]. The device generates a random number between 0 and 1 and compares it with the ACB factor broadcasted by the BS. The device can only access the BS if the generated number is less than the ACB factor. Although the literature has a vast amount of works extending the basic idea of the ACB, such as extended ACB [10], cooperative ACB [11], and dynamic ACB [12], ACB still fails to satisfy strict latency requirements [13].

To this end, the need of extreme low latency in IoT urges the design of novel access schemes to overcome the flaws associated with the old ones. Learning-based schemes were discussed in many surveys as the potential solution to the existing problems of the proposed approaches in the literature to overcome the RA limitations [14]. Moreover, many emerging IoT applications can exploit activation correlation and

traffic prediction to enable pre-emptive resource allocation and achieve ultra-reliable and low latency communications. The traffic correlation behavior of MTDs enables traffic prediction and forecasting algorithms to anticipate the set of active and silent MTDs. In this context, Fast Uplink (FU) grant was introduced in [15] to allow for resource allocation based on traffic prediction schemes.

### A. Fast Uplink Grant

To elaborate more on FU grant, we consider  $K$  IoT devices and  $L$  available transmission slots, where  $K \gg L$ . Each device is stimulated to generate data packets at different time slots controlled by different processes at the application layer, e.g., triggered external events. Whenever a device generates a packet, it will need a transmission slot to transmit it to the BS. In the FU scheme, the BS allocates the available transmission slots to the set of IoT devices that it believes will transmit in the current time slot. The designed resource allocation scheme should exploit the correlation of traffic pattern based on the temporal and event dimensions.

The FU scenario relies mainly on traffic prediction. The BS has to efficiently predict the probability of each device to be active or silent and grant the available resources to those most likely to be active, with some fairness guarantees. Some of the potential advantages of applying FU are:

- Absence of scheduling requests and collisions leading to a reduction in the energy consumption of IoT devices and uplink latency;
- Clearance of signaling overhead between the devices since learning occurs only at the side of the BS;
- It allows for the potential use of the uplink grant signal to partially or fully estimate the channel condition at the IoT devices side before actual uplink process (CSIT)<sup>1</sup>.

### B. Contributions

In this work, we build upon [16], where we define the main system model that consists of a set of binary discrete events that affects the activation patterns of massive IoT devices. The binary events are modeled as Markovian sources. We introduce an FU algorithm that exploits the traffic correlation to efficiently predict the IoT devices' traffic pattern using hidden Markov model (HMM) and the forward algorithm. The forward algorithm is a learning algorithm that fits the proposed HMM. The results show that the FU algorithm outperforms the conventional RA and TDMA schemes in terms of the accuracy and efficiency of resource allocation.

Another novel contribution is that we post-process the prediction of the forward algorithm to lower the average experienced for AoI all devices at each time step while maintaining the prediction accuracy as high as possible. We optimize an age parameter to increase the resulting allocation index of the high-age devices and guarantee a higher degree of scheduling fairness. In addition, we formulate a

<sup>1</sup>Notice that channel estimation is a proposed advantage when applying the FU scheme, e.g., via pilot symbols transmitted within the uplink grant signal. However, we leave this work for future implementation.

baseline model based on the forward algorithm that forms a distribution of the activation probability using extremely low computation resources. Furthermore, we estimate the model hyperparameters to exploit the formulated FU algorithm in real-time applications without prior knowledge of the model hyperparameters. We then propose an online-learning version of the FU algorithm, where the BS exploits only the set of observations at each instant to allocate the resources to the devices using the learnt hyperparameters. The simulation results illustrate that applying the online-learning algorithm at each instant still captures the age and the accuracy of the actual genie-aided model and outperforms the traditional resource allocation schemes and the HMM baseline scheme.

The contributions of this work are summarized as follows:

- We formulate the device activation probabilities for the described HMM system model.
- We apply the forward algorithm to predict the active devices and perform preemptive FU grant with low complexity.
- We optimize an age parameter to compensate the AoI of the devices that have experienced high AoI while preserving the accuracy of the efficient forward algorithm.
- For the case of unknown hyperparameters of the model, we apply an expectation-maximization algorithm to estimate the event transition probabilities and the device activation probabilities based only on the observations. Then, we apply the estimation procedure to present an offline-learning version of the FU algorithm.
- Finally, we rely on both the AoI compensation and the learned parameters to formulate an online-learning scheme that allows the BS to perform the FU algorithm in real-time applications, without prior availability of large activation data sets.
- The proposed online and offline schemes clearly outperform conventional GF and TDMA in terms of resource allocation efficiency while guaranteeing a favourable amount of fairness via age compensation.

### C. Outline

The rest of the paper is organized as follows: Section II discusses the related literature. Section III depicts the system model for the IoT device. It also explains performance metrics that are used to evaluate the performance of the proposed FU schemes. Next, Section IV applies the forward algorithm to predict the traffic pattern of IoT devices. After that, Section V discusses the online-learning version of the FU algorithm. Section VI depicts and discusses different results for the performance evaluation. Finally, Section VII concludes the paper and discusses future research directions.

**Notation:** Boldface lowercase letters denote vectors.  $Pr$  denotes the probability equation. In addition,  $[x]^+$  refers to  $\max(0, x)$ ,  $\arg \max$  is the maximization notation, and  $\arg \min$  is the minimization notation.  $\bar{x}$  is the mean of  $x$  and  $C(a, b)$  is the cost function, where  $a$  and  $b$  are the parameters to be optimized. To make the paper more tractable, we summarize the key abbreviations and symbols that will appear throughout the paper in Table I.

TABLE I: Important abbreviations and symbols.

ACB	access class barring
AoI	age-of-information
CMAP	coupled Markovian arrival process
CMMPP	coupled Markov modulated Poisson process
DRL	deep reinforcement learning
FU	fast uplink
GF	grant-free
HMM	hidden Markov model
LSTM	long short-term memory
MTD	machine-type communication device
NOMA	non-orthogonal multiple access
PDF	probability density distribution
RA	random access
RNN	recurrent neural network
SVM	support vector machine
TDMA	time division multiple access
$K$	number of IoT devices
$L$	number of frequency resources
$N$	number of Markovian events
$Z$	number of Baum-Welsh iterations
$\epsilon_0, \epsilon_1$	temporal transition probabilities
$q_{nk}$	activation probabilities
$\omega_t$	wrong allocations
$\mu_t$	missed allocations
$\beta$	age parameter
$I_{t+1}^{(k)}$	scheduling priority index
$\bar{R}$	average regret
$\bar{\Delta}$	average age

## II. STATE OF THE ART

Many learning-based schemes have been proposed in the literature for resource allocation in IoT networks. In this section, we present a brief literature review of the existing schemes and discuss their limitations. To begin with, in [17], [18], the authors studied the activation of devices following coupled Markov modulated Poisson process (CMMPP) and coupled Markovian arrival process (CMAP) traffic models, respectively. However, they did not offer resource allocation schemes based on these traffic models. In [19], the authors used an HMM model to build a decision fusion algorithm that investigates the correlation time between binary sources in a wireless sensor network (WSN). In the same context, the work in [20] exploited the correlated activity of devices to develop heuristic protocols for GF RA. Sinusoidal spreading sequences were proposed in [21] to enable FU grant based on free non-orthogonal multiple access (NOMA), whereas authors in [22] discussed hybrid resource allocation schemes to overcome the large signaling overhead and collision problems resulting from message replications in GF transmission. Moreover, in [23], Samad et al. introduced a multi-armed bandit algorithm to perform FU grant in IoT networks. However, this work also came short from exploiting the traffic correlation on the event-temporal basis.

The authors in [24] present an FU grant algorithm based on support vector machines (SVM) and long short-term memory (LSTM). However, the addressed algorithm needs efficient hardware at the BS to carry out complex neural networks computations. Authors in [25] presented an FU grant-based federated learning approach, where the BS relies on the traffic estimation at the side of the devices. Although performing the estimation at the side of the devices side reduces the complexity at the BS side, which is responsible only to perform allocation, it requires the low power end devices to perform complex computations. In addition, authors in [26] formulate a reinforcement learning algorithm for resource allocation in device-to-device (D2D) communications, whereas authors in [27] propose a recurrent neural network (RNN) model based on meta-learning to predict the millimeter wave (mmWave) link blockages. Mohammadi et al. [28] presented a multi-agent deep reinforcement learning (DRL) solution for resource allocation, authors in [29] proposed a clustering-based solution to perform resource scheduling depending on each cluster priority and demands, and the work in [30] presented a survey of recent artificial intelligent (AI)-based frameworks for resource allocation in diverse use cases. Table II summarizes the existing reviewed literature.

The majority of the referred literature relies on the use of machine learning and reinforcement learning schemes, which need to perform complex computations either at the BS side or the IoT devices side. This requires powerful hardware and a long training duration that reflects some challenges on the usage of machine learning in communication systems [31]. In addition, IoT networks are often driven by interactive applications, where observations are provided based on human/machine interaction over time, which means that adding a set of new observations to the collected observations for a period of time changes the model and the learning problem. Therefore, online learning becomes necessary [32]. Hence, generalized complex machine learning schemes might not be able to train real-time IoT networks as they require extremely powerful hardware to perform their learning algorithms online and simpler, specially tailored, learning schemes are required for online learning scheduling algorithms [33]. In this work, we present a stochastic-based solution, which fits well with the proposed HMM model. Moreover, it is very efficient in terms of prediction accuracy and simpler than the existing machine learning solutions in the literature.

### III. SYSTEM MODEL AND PROBLEM FORMULATION

Consider an IoT network, such as NB-IoT, with  $K$  IoT devices relay their information to a single BS as depicted in Fig. 2. As in conventional LTE FU, the transmission resources are link is divided into time slots, and in every time slot the BS can schedule up to  $L$  devices for transmission in  $L$  frequency slots. The scheduled devices are assigned to transmission slots and they transmit only if they are active (i.e. if they have data to transmit). If a device is scheduled for transmission while inactive, the uplink resource is wasted.

We denote the activation of device  $k$  in discrete time slots  $t = 1, 2, \dots$  by the random variable  $A_t^{(k)}$ .  $A_t^{(k)} = 1$  if

TABLE II: Summary of the literature review.

Main scope	Literature	Sub-topics
Traffic Models	M. Laner et al. [17]	CMMPP
	E. Grigoreva et al. [18]	CMAP
	P. Rossi et al. [19]	HMM in WSN
RA-based	A. E. Kalør et al. [20]	GF-RA
	S. M. Hasan et al. [21]	NOMA FU grant
	Z. Zhou et al. [22]	Hybrid resource allocation
FU grant	S. Ali et al. [23]	Multi-armed bandit
	E. Eldeeb et al. [24]	SVM and LSTM
	O. Habachi et al. [25]	Federated learning
Deep learning	I. AlQerm et al. [26]	Reinforcement learning
	A. E. Kalør et al. [27]	Meta-learning and RNN
	F. Mohammadi et al. [28]	Multi-agent DRL
	X. Liu et al. [29]	Clustering
	D. Hejji et al. [30]	AI survey

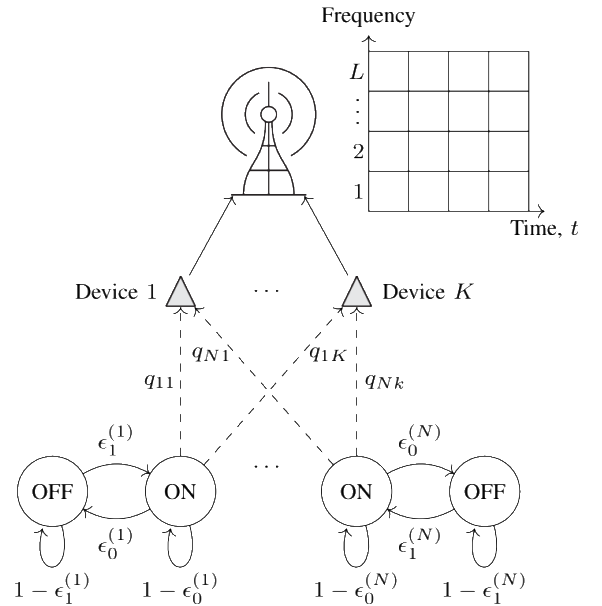


Fig. 2: The considered activation model, in which  $N$  On-Off Markovian processes control the activation of  $K$  devices. If process  $n$  is in the On-state it activates device  $k$  with probability  $q_{nk}$ .

the device is active, otherwise  $A_t^{(k)} = 0$ . The activation of IoT devices at time  $t$  is indicated by the vector  $\mathbf{A}_t = \{A_t^{(1)}, \dots, A_t^{(K)}\}$ .

#### A. State Transition Probabilities

The activation of the devices is controlled by  $N$  independent two-state Markov processes. The Markovian processes swing between On and Off states, where at time  $t$ , the state  $\mathcal{S}_t^{(n)} \in \{1, 0\}$ , is governed by temporal transition probabilities  $\epsilon_1^{(n)}$ ,  $\epsilon_0^{(n)}$  as shown in Fig. 2, where

$$\Pr(\mathcal{S}_{t+1}^{(n)} = 0 | \mathcal{S}_t^{(n)} = 1) = \epsilon_0^{(n)}, \quad (1)$$

$$\Pr(\mathcal{S}_{t+1}^{(n)} = 1 | \mathcal{S}_t^{(n)} = 0) = \epsilon_1^{(n)}, \quad (2)$$

$$\Pr(\mathcal{S}_{t+1}^{(n)} = 0 | \mathcal{S}_t^{(n)} = 0) = 1 - \epsilon_1^{(n)}, \quad (3)$$

$$\Pr(\mathcal{S}_{t+1}^{(n)} = 1 | \mathcal{S}_t^{(n)} = 1) = 1 - \epsilon_0^{(n)}. \quad (4)$$

To this end, we define the state vector at time  $t$  as  $\mathbf{S}_t = \{\mathcal{S}_t^{(1)}, \dots, \mathcal{S}_t^{(N)}\}$ . The Markov processes that are in the On state, i.e.  $\mathcal{S}_t^{(n)} = 1$ , may activate specific IoT devices, where the probability that Markov process  $n$  activates device  $k$  is given by  $q_{nk}$ .

### B. Device Activation Probabilities

A certain device becomes active if one or more of the Markovian states activates it. Thus, the probability that device  $k$  is active at time  $t$  is

$$\Pr\left(A_t^{(k)} = 1 \mid \mathbf{S}_t\right) = 1 - \prod_{n=1}^N \Pr\left(A_t^{(k)} = 0 \mid \mathcal{S}_t^{(n)}\right) \quad (5)$$

$$= 1 - \prod_{n=1}^N (1 - q_{nk})^{\mathcal{S}_t^{(n)}}, \quad (6)$$

where the activation is considered to be conditionally independent given the state vector  $\mathbf{S}_t$ .

Furthermore, the probability that IoT device  $k$  will be active at the future time instant  $t + 1$  given the state vector at time  $t$  can be written as

$$\Pr\left(A_{t+1}^{(k)} = 1 \mid \mathbf{S}_t\right) = 1 - \prod_{n=1}^N \Pr\left(A_{t+1}^{(k)} = 0 \mid \mathcal{S}_t^{(n)}\right) \quad (7)$$

$$= 1 - \prod_{n=1}^N h(n), \quad (8)$$

where

$$h(n) = \begin{cases} 1 - \epsilon_1^{(n)} + \epsilon_1^{(n)}(1 - q_{nk}), & \mathcal{S}_t^{(n)} = 0 \\ \epsilon_0^{(n)} + (1 - \epsilon_0^{(n)})(1 - q_{nk}), & \mathcal{S}_t^{(n)} = 1. \end{cases} \quad (9)$$

### C. Performance Evaluation Metrics

Next, we define key performance metrics that are essential to evaluate the proposed FU scheme with traffic prediction and compare it to existing allocation schemes.

1) *Regret*: The *regret* is one of the key metrics used to evaluate the performance of scheduling algorithms using learning schemes [15]. We define one unit of regret as wasting a resource on an inactive device while one active device did not receive a resource. Therefore, regret is the accumulated regret units at each time slot that resulted from the prediction and scheduling of active devices. Consider the uplink grant vector  $\mathbf{U}_t = \{u_t^{(1)}, \dots, u_t^{(K)}\}$ , where  $u_t^{(k)} = 1$  if a slot is allocated to device  $k$  at time  $t$  and  $u_t^{(k)} = 0$  if device  $k$  does not receive a transmission slot. The number of wrong allocations at time instant  $t$  can be calculated as the difference between the uplink grant vector  $u_t^{(k)}$  at time instant  $t$  and the activation vector  $A_t^{(k)}$  at time instant  $t$  as follows

$$\omega_t = \sum_{k=1}^K \left[ u_t^{(k)} - A_t^{(k)} \right]^+, \quad (10)$$

where  $[x]^+ = \max(0, x)$ . In addition, the number of missed allocations can be computed as the difference between the

activation vector  $A_t^{(k)}$  at time instant  $t$  and the uplink grant vector  $u_t^{(k)}$  at time instant  $t$  as follows:

$$\mu_t = \sum_{k=1}^K \left[ A_t^{(k)} - u_t^{(k)} \right]^+. \quad (11)$$

Hence, the regret function at time  $t$  is defined as

$$R(t) = \min\{\omega_t, \mu_t\}. \quad (12)$$

Then, minimizing the long-term  $R(t)$  is an important target, when designing an FU grant scheme.

The meaning of the regret function can be understood by considering the following three cases. First, if  $M > L$  devices are active and all the  $L$  uplink grants are given to a subset of the active devices, then  $\omega_t = 0$  and  $\mu_t = 0$ . This results in a regret of  $R(t) = 0$ , reflecting that the number of unserved devices is minimized. If no devices are active, and the  $L$  grants are given to inactive devices,  $\omega_t = L$  and  $\mu_t = 0$ . This also results in  $R(t) = 0$ , again reflecting a minimum number of unserved devices. Finally, if  $M \leq 2L$  devices and scheduler assigns grants to  $M/2$  of the active devices and  $L - M/2$  inactive devices, then  $\omega_t = L - M/2$  and  $\mu_t = M/2$ . The regret is then  $R(t) = \min(L - M/2, M/2)$ , which renders the number of unserved devices that could have been served if the allocation process was more accurate.

2) *System usage*: We propose the system usage metric which would help with evaluation the efficiency of the proposed FU grant allocation scheme. The average system usage  $\eta_t$  at time  $t$  is defined as the ratio between the number of transmission slots that are successfully used by an IoT device to the total number of available slots  $L$  averaged over time. That is

$$\eta_t = \frac{1}{tL} \sum_{\tau=0}^t L - \omega_\tau. \quad (13)$$

The average system usage marks the percentage of transmission slots that are successfully used for uplink by the IoT devices.

3) *Age of Information*: To measure the freshness of data and the degree of fairness in scheduling the devices, we define the discrete AoI [5], [34] of device  $k$  as the time passed since the device transmitted a packet. That is the last time instant in which device  $k$  was active and received a transmission grant and

$$\Delta^{(k)} = t - t_k, \quad (14)$$

where  $t_k < t$  is the last time slot before  $t$ , when  $A_{t_k}^{(k)} = u_{t_k}^{(k)} = 1$  and the AoI should be a non-negative integer. The average age per device at a certain time is defined as

$$\bar{\Delta} = \frac{1}{K} \sum_{k=1}^K \Delta^{(k)}. \quad (15)$$

Meanwhile, the peak age per device can be noted as  $\max_k \{\Delta^{(k)}\}$ .

AoI is important in the proposed scenario since it provides a measure for the freshness of the data received from each IoT device. This means that if a device is rarely scheduled for transmission, the information stored at the BS from this device

will be outdated as the device's age becomes too high. Hence, it is also considered as a measure of fairness, where higher average ages mean that some devices are rarely scheduled and low average age means that devices are fairly scheduled.

**Remark 1.** We assume that the BS has pre-knowledge of the environment, and hence knows the state transition probabilities  $\epsilon^{(n)}$  and the device activation probabilities  $q_{nk}$ . Therefore, the BS aims to jointly minimize the regret and the AoI and maximize the system usage by scheduling the available transmission resources to the devices. In addition, we investigate the same objective while assuming that the state transition probabilities and the device activation probabilities are not fully known by the BS. Hence, the BS needs to estimate the model hyperparameters via estimation algorithms.

#### IV. THE PROPOSED FAST UPLINK ALGORITHM

This section analyzes the device's temporal activation probabilities and exploits them to develop the traffic prediction-based FU scheme. The BS uses the set of past observations of each device to predict the hidden states for each event. Afterward, it uses the set of predicted hidden states to generate an estimate for the future observations for each device.

##### A. Traffic Prediction

The BS does not know the states of the Markov processes and hence, continuously needs to estimate them based on the observations. Notice that the activation process of the IoT devices can be described by an  $N$ -HMM as typically detailed in [35]. Concretely, the forward algorithm can be applied by the BS to learn the probability of events being in a certain state given the history of IoT devices activation observations done by the BS [36]. The BS can exploit the learned state distribution to estimate future device activation probabilities and patterns.

To obtain a clear understanding of the forward algorithm, consider the joint probability  $p(\mathbf{S}_t, \mathbf{A}_t)$ . The forward algorithm is able to efficiently compute this joint probability in a recursive way as in [37]. Herein, the forward algorithm is described as follows

$$p(\mathbf{S}_t, \mathbf{A}_{1:t}) = p(\mathbf{A}_t | \mathbf{S}_t) \sum_{\mathbf{S}_{t-1}} p(\mathbf{S}_t | \mathbf{S}_{t-1}) p(\mathbf{S}_{t-1}, \mathbf{A}_{1:t-1}). \quad (16)$$

Then the most likely hidden state for the events can be learned using

$$\mathbf{S}_t^* = \arg \max_{\mathbf{S}_t} p(\mathbf{S}_t, \mathbf{A}_{1:t}). \quad (17)$$

The estimated hidden states at time instant  $t$  are used to predict the activation probabilities of each device at time instant  $t+1$  using (7). The predicted device activation probabilities can be formulated as

$$\Pr(A_{t+1}^{*(k)} = 1 | \mathbf{S}_t^*) = 1 - \prod_{n=1}^N \Pr(A_{t+1}^{(k)} = 0 | \mathcal{S}_t^{*(n)}). \quad (18)$$

Alternatively, the BS can use the forward algorithm results directly to predict the maximum likelihood of the pattern of the devices in the next time instant

$$\begin{aligned} \mathbf{A}_{t+1}^* &= \arg \max_{\mathbf{A}_{t+1}} \sum_{\mathbf{S}_t} \Pr(\mathbf{A}_{t+1} | \mathbf{S}_t) p(\mathbf{S}_t, \mathbf{A}_{1:t}) \\ &= \arg \max_{\mathbf{A}_{t+1}} \sum_{\mathbf{S}_t} p(\mathbf{S}_t, \mathbf{A}_{1:t}) \prod_{k=1}^K \Pr(A_{t+1}^{(k)} = b_k | \mathbf{S}_t), \end{aligned} \quad (19)$$

$$(20)$$

where  $\mathbf{A}_{t+1}^*$  is the maximum likelihood estimate of the set of active IoT devices at time  $t+1$ , and  $b_k \in \{1, 0\}$ .

Note that (19) evaluates the probability of a full pattern. Hence, it gives the most likely activation pattern and does not consider the activation probability of each device separately. Meanwhile, when performing uplink grant allocation, the BS should select the  $L$  devices which are most likely to be *jointly* active. In order to determine these devices, we assume that the system is in the most likely state, found from (17), and exploit this assumption to compute the transition probability of the events as follows

$$P_{\text{On}}^{(n)} = \Pr(\mathcal{S}_{t+1}^{*(n)} = 1 | \mathcal{S}_t^{*(n)} = 1), \quad (21)$$

which will be used to determine the activation likelihood of each device as

$$P_{\text{device}}^{(k)} = P_{\text{On}}^{(1)} \cdot q_{1,k} \cup \quad (22)$$

$$P_{\text{On}}^{(2)} \cdot q_{2,k} \cup \quad (23)$$

$$\vdots \cup \quad (24)$$

$$P_{\text{On}}^{(n)} \cdot q_{n,k}. \quad (25)$$

Finally, the devices are sorted by their activation probability, and the  $L$  devices most likely to be active are scheduled in the next slot.

##### B. Baseline Model

We develop a baseline model that can capture the behavior of the devices efficiently with low computational complexity using the steady-state probabilities of the events  $p(S_{t_{ss}}^{(n)})$  as follows

$$\Pr(A_{t+1}^{(k)} = 1) = 1 - \sum_{\mathbf{S}_t} \prod_{n=1}^N (1 - q_{nk})^{\mathcal{S}_t^{(n)}} p(S_{t_{ss}}^{(n)}), \quad (26)$$

where

$$p(S_{t_{ss}}^{(n)}) = \begin{cases} \frac{\epsilon_0^{(n)}}{\epsilon_0^{(n)} + \epsilon_1^{(n)}}, & \mathcal{S}_t^{(n)} = 0, \\ \frac{\epsilon_1^{(n)}}{\epsilon_0^{(n)} + \epsilon_1^{(n)}}, & \mathcal{S}_t^{(n)} = 1. \end{cases} \quad (27)$$

The steady-state probabilities of the events, as calculated in (27), describe how likely each state will be active long enough during the simulation time [38]. We formulate a probability density distribution (PDF) by multiplying the steady-state probabilities the device activation probabilities as in (26). This PDF describes the probability of a device to be active affected by the steady-state probability of the states. Hence,

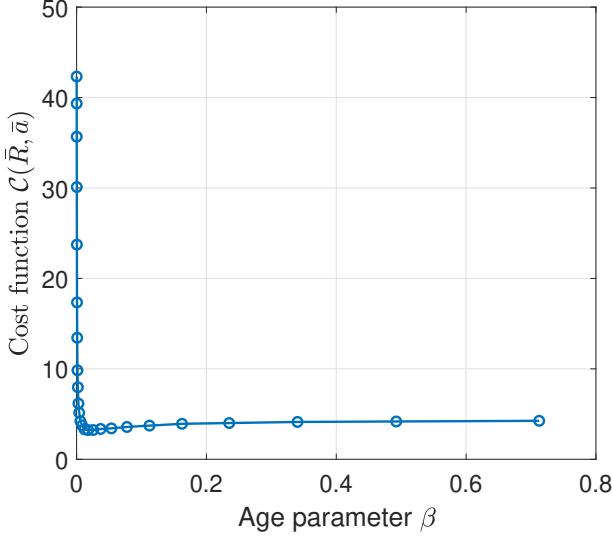


Fig. 3: Cost function  $\mathcal{C}(\bar{R}, \bar{\Delta})$  for the age parameter  $\beta$  optimization.

this distribution gives a simple description of the activation pattern of the devices without performing any forecasting computations. Afterward, the devices are scheduled by the BS according to this distribution. Note that we refer to this scheduling algorithm as the baseline model.

### C. AoI Compensation

We introduce the age parameter  $\beta$  to map the priority of scheduling devices that have high AoI. Higher values of  $\beta$  mean that the BS gives higher priority to devices that have not transmitted for a long time (i.e., devices with higher AoI). The scheduling priority index for device  $k$  at time  $t+1$  is thus defined as

$$I_{t+1}^{(k)} = \Pr(A_{t+1}^{(k)} = 1 | \mathbf{S}_t) + \beta p(S_{t_{ss}}) \Delta^{(k)}(t) \quad (28)$$

$$= 1 - \prod_{n=1}^N h(n) + \beta p(S_{t_{ss}}) \Delta^{(k)}(t). \quad (29)$$

Instead of sorting the devices according to their probability of activation, the BS sorts the devices according to their index  $I$ . Then the  $L$  devices with the highest index  $I$  are scheduled for transmission.

The BS needs to choose an appropriate value for  $\beta$  in (29) to control the trade-off between the devices' AoI and regret optimalities. This introduces an optimization problem at the BS side, where the cost function  $\mathcal{C}(\bar{R}, \bar{\Delta})$  is defined as the multiplication of the average regret  $\bar{R}$  and the average AoI  $\bar{\Delta}$

$$\arg \min_{\beta} \mathcal{C} = \bar{R} \cdot \bar{\Delta}, \quad (30)$$

$$s.t. \quad \beta \geq 0. \quad (31)$$

As illustrated in Fig. 3, we can notice that the cost function is convex and can be optimized easily to get the optimal  $\beta$  that lowers down the AoI while maintaining the regret in an appropriate region for a given network setup.

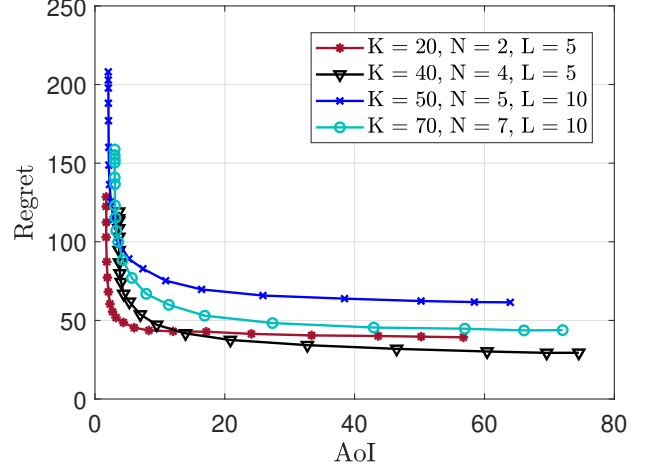


Fig. 4: Achievable region for the AoI and the regret while applying the AoI compensation for different values of the age parameter  $\beta$ .

To address the trade-off between the AoI and the regret, we investigate Fig. 4 that depicts the achievable region for AoI and regret using different values of  $\beta$  for different setup of the network (the number of devices, the number of binary events, and the available number of resources). The smaller the network setup  $K, N$ , and  $L$ , the smaller the resulting age and regret. Therefore, each BS needs to optimize its own  $\beta$  according to the prior knowledge of the network parameters. If  $\beta$  is set to 0, the scheduling resets to its basic form without the age compensation term (fair regret), whereas if  $\beta$  is set to asymptotically  $\infty$ , the scheduler will act as round-robin, where the resources are distributed equally among the devices (fair age).

### V. ONLINE LEARNING BASED ON MODEL ESTIMATION

The forward algorithm and the HMM mainly depend on prior knowledge of the hyperparameters of the model, namely, the transition state probabilities for each event and the activation probabilities when affected by active events. Sometimes, it is difficult to have prior knowledge of these parameters. Therefore, the BS aims at estimating the hyperparameters of the model using only the possible observations from the real-time model. Next, we present the estimation algorithm for both  $q_{nk}$  and  $\epsilon$ .

The activation probabilities of device  $k$  at time instant  $t$  given the set of states  $\mathbf{S}_t$  are the set of values that result in an activation pattern that is as close as possible to the actually observed activation pattern  $A_t^{(k)}$ . To estimate  $q_{nk}$ , we formulate the following likelihood maximization formula

$$\mathbf{q}_{nk}^* = \arg \max_{\mathbf{q}_{nk}} \prod_{t=1}^T \Pr(A_t^{(k)} = b_k | \mathbf{S}_t^*), \quad (32)$$

where  $b_k \in \{1, 0\}$ ,  $\Pr(A_t^{(k)} = b_k | \mathbf{S}_t^*)$  is calculated as follows

$$\Pr(A_t^{(k)} = b_k | \mathbf{S}_t^*) = \begin{cases} 1 - \prod_{n=1}^N (1 - q_{nk}) S_t^{*(n)}, & b_k = 1, \\ \prod_{n=1}^N (1 - q_{nk}) S_t^{*(n)}, & b_k = 0, \end{cases} \quad (33)$$



with the constraint  $0 < q_{nk} < 1$ . Note that (32) can be solved via geometric programming which can be solved for each device  $k$  using any programming tool, such as fmincon, which is available in Matlab, or cvx (available in both Matlab and Python) [39], or even using a basic exhaustive search algorithm to find the solution of the optimization problem. In this context, the cvx tool is considered the best fit for such complex problems with multiple local maxima, where it can solve geometric programming problems efficiently. However, the optimization problem relies on predicting the most likely hidden state  $S_t^*$  from (17) using the forward algorithm, which uses the actual hyperparameter values  $q_{nk}$  and  $\epsilon$ . This problem can be solved iteratively using the Baum-Welsh algorithm [37].

The Baum-Welsh method relies on the forward-backward algorithms, where at time instant  $t$ , it estimates the expected number of visits of each state and the number of transitions from state  $S_i$  to state  $S_j$  during the time period  $T$  ( $0 \leq T \leq t$ ). Afterward, it exploits the number of visits and transitions to generate an estimate of  $\epsilon^*$ . The estimated temporal transition probabilities  $\epsilon^*$  along with the previous estimate of  $q_{nk}^*$  are used to predict the most likely hidden state, which will be used to update the estimate of  $q_{nk}^*$ . These iterations are repeated until convergence (desired error threshold). It is expected that the Baum-Welsh algorithm<sup>2</sup> converges after a limited number of iterations  $Z$  according to the complexity of the model. After convergence, we can exploit the estimated hyperparameter values  $q_{nk}^*$  and  $\epsilon^*$  to perform resource allocation for the devices. After initializing  $q_{nk}(0)$ ,  $\epsilon_0^{(n)}(0)$  and  $\epsilon_1^{(n)}(0)$ , we apply the following equations that illustrate the expectation-maximization estimation procedure

$$\mathbf{S}_t^*(i) = \arg \max_{\mathbf{S}_t} p(\mathbf{S}_t, \mathbf{A}_{1:t}) \Big|_{q_{nk}=q_{nk}^*(i-1), \epsilon_{b_k}^{(n)}=\epsilon_{b_k}^{*(n)}(i-1)}, \quad (34)$$

$$\Pr \left( A_t^{(k)} = b_k \mid \mathbf{S}_t^*(i) \right) = \begin{cases} 1 - \prod_{n=1}^N (1 - q_{nk}) S_t^{*(n)}(i), & b_k = 1, \\ \prod_{n=1}^N (1 - q_{nk}) S_t^{*(n)}(i), & b_k = 0, \end{cases} \quad (35)$$

$$\mathbf{q}_{nk}^*(i) = \arg \max_{\mathbf{q}_{nk}} \prod_{t=1}^T \Pr \left( A_t^{(k)} = b_k \mid \mathbf{S}_t^*(i) \right). \quad (36)$$

In fact, this learning process requires enough number of observations to ensure an accurate estimation procedure. If the BS has prior knowledge to a number of observations that is large enough to perform the estimation, we refer to it as FU-offline learning. On the other hand, applying this iterative expectation-maximization procedure at each time-step converts the ordinary algorithm to an online version of the FU algorithm. First, the BS collects the observations at time instant  $t$ , where it utilizes them to iteratively estimate the model hyperparameters  $\mathbf{q}_{nk}$  and  $\epsilon$ . Afterward, it predicts the activation pattern probability of each device at time instant  $t+1$  using the forward algorithm. Moreover, it optimizes the age parameter  $\beta$  to compensate for the age of the devices that

<sup>2</sup>A more interested reader can refer to [37] for more details about the Baum-Welsh expectation-maximization algorithm.

---

**Algorithm 1:** Traffic prediction based fast uplink grant algorithm.

---

```

1  $t = 1$ .
2 Define  $K, N, L$ , and  $Z$ .
3 Initialize the age vectors  $\Delta^{(k)}$ .
4 Initialize the regret vectors  $R^{(k)}$ .
5 while True do
6   Initialize  $q_{nk}(0)$ ,  $\epsilon_0^{(n)}(0)$  and  $\epsilon_1^{(n)}(0)$ .
7   Collect the observations  $A_t$ .
8   for  $i = 1, \dots, Z$  do
9      $\mathbf{S}_t^*(i) =$ 
        $\arg \max_{\mathbf{S}_t} p(\mathbf{S}_t, \mathbf{A}_{1:t}) \Big|_{q_{nk}=q_{nk}^*(i-1), \epsilon_{b_k}^{(n)}=\epsilon_{b_k}^{*(n)}(i-1)}$ .
10    Update  $\epsilon_{b_k}^{*(n)}(i)$ .
11     $\mathbf{q}_{nk}^*(i) =$ 
        $\arg \max_{\mathbf{q}_{nk}} \prod_{t=1}^T \Pr \left( A_t^{(k)} = b_k \mid \mathbf{S}_t^*(i) \right)$ .
12  end
13  Optimize the age parameter  $\beta$ .
14  Compensate  $\Pr \left( A_{t+1}^{(k)} = 1 \mid \mathbf{S}_t^* \right)$  using  $\beta$ .
        $I_{t+1}^{(k)} = \Pr \left( A_{t+1}^{(k)} = 1 \mid \mathbf{S}_t \right) + \beta p(S_{t_{ss}}) \Delta^{(k)}(t)$ .
15  Allocate the  $L$  resources.
16  Update the age vector  $\Delta^{(k)}$  for each device.
17  Update the regret vector  $R^{(k)}$  for each device.
18   $t = t+1$ .
19 end

```

---

experience high age. Finally, the BS allocates the resources to the devices with the highest priority index. We refer to this procedure as online learning-enhanced AoI, which is depicted in Algorithm 1.

## VI. RESULTS AND DISCUSSION

In this section, we present the simulation results of the proposed FU algorithm based on the forward algorithm and the further discussed extensions. We consider a setup of a single BS with  $L = 10$  available frequency resources at each time instant and  $K = 50$  sensors affected by  $N = 5$  Markovian events. The temporal state transition probabilities are  $\epsilon_0^{(n)}$  and  $\epsilon_1^{(n)}$  are uniformly distributed on the interval  $[0, 0.5]$ . Note that, low values of  $\epsilon$  result in forcing the events to be active for longer times and cause congested traffic. Meanwhile the activation probabilities  $q_{nk} \in [0, 1]$ . We present a detailed comparison between the proposed algorithms and some of the existing models. For instance, we discuss the GF, where the active devices send a request to the BS using a random preamble, and the TDMA, where round-robin is followed to schedule the resources for the devices. In addition, we present the FU-genie-aided that refers to the case in which the states of the events are assumed to be perfectly known to the BS. Herein, the FU-limited info refers to the scenario in which the BS observes only the activation of the scheduled sensors. Meanwhile, in the FU-feedback, the BS is allowed to also observe the activation of the devices that were not scheduled



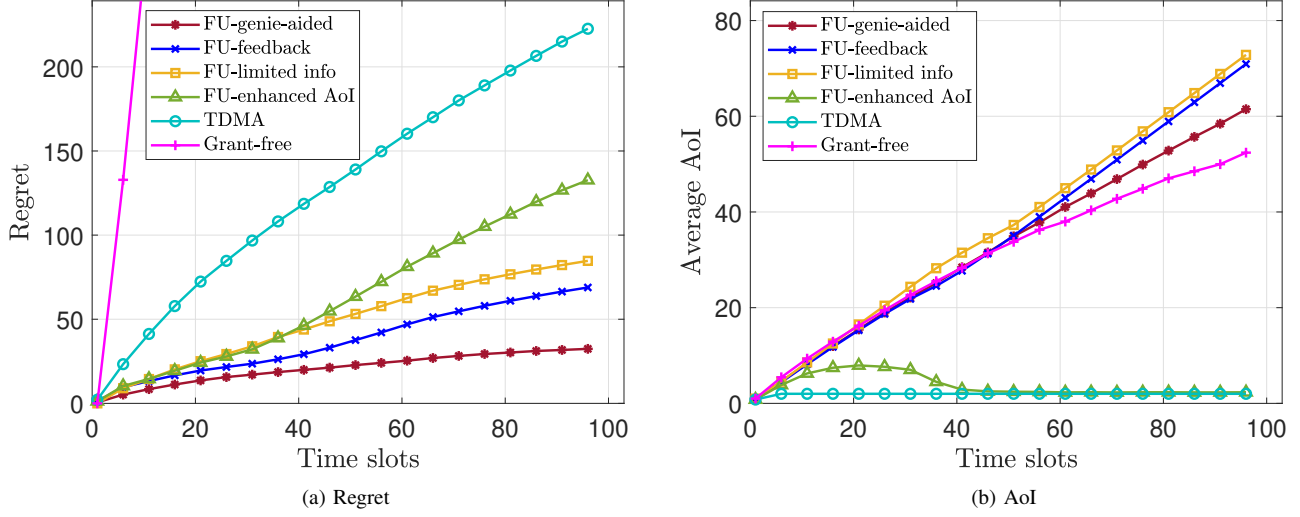


Fig. 5: Regret and AoI evaluation.  $K = 50$  sensors,  $N = 5$  events, and  $L = 10$  available frequency resources.

TABLE III: The parameters used in the simulation setup.

Parameter	Value	Parameter	Value
$K$	50	$N$	5
$L$	10	$T$	100
$Z$	40	$\beta$	0.0233
$\epsilon_0^n, \epsilon_1^n$	$[0, 0.5]$	$q_{nk}$	$[0, 1]$

through a feedback signal. The FU-baseline is presented as the low computational version of the FU algorithm as presented in IV-B. The term FU-enhanced AoI corresponds to the FU algorithm after performing the age compensation as discussed in IV-C. Finally, FU-offline learning corresponds to applying the estimation algorithm discussed in V while assuming a prior knowledge of enough observations offline to be used to estimate the model hyperparameters, whereas online learning-enhanced AoI is the online version of the presented algorithm, where no prior information is assumed to be known and age compensation is applied as discussed in algorithm 1. Table III illustrates the parameters used in the simulation.

Fig. 5 demonstrates the regret and the average AoI performance metrics when applying the discussed schedulers. In Fig. 5-(a), we evaluate the regret function, where the FU-feedback scheme significantly outperforms both GF and TDMA. Specifically, when applying the proposed FU-feedback scheme, the regret function is reduced to 4 times less than the regret in the case of TDMA and 50 times less than the regret of GF due to the high number of collisions in GF. Moreover, the FU-limited info scheme has close results in terms of the regret to the genie-aided model which assumes perfect knowledge of the events. The feedback version of the FU algorithm exploits the cost of having imperfect information about the activation of the devices, which reflects on the resulting regret. However, the performance is still close to that of the genie-aided model and outperforms existing models (GF and TDMA).

Fig. 5-(b) shows the average AoI per device, where the

proposed FU-feedback scheme has relatively higher ages when compared to GF and TDMA, which motivates the need for an enhanced AoI version of the FU algorithm. In addition, we calculate the system usage using (13), where the FU-feedback achieves nearly a 0.95 system usage, which indicates that the BS has successfully allocated 95% of the resource to the transmitting devices. Hence, the proposed scheme is more efficient than TDMA which uses only 78% of the resources, and the GF that has only 50% of system usage due to the high number of collisions.

Solving the optimization problem in (31) renders  $\beta = 0.0233$  as the optimal value for the addressed setup. The BS applies the age parameter  $\beta$  to address the fairness issue. Fig. 5 shows the age enhancement which results from applying the fairness parameter  $\beta = 0.0233$  while scheduling the devices. The average age per device for the FU-enhanced AoI is significantly improved when compared to the basic implementation with  $\beta = 0$ . The average age per device is much lower than GF and asymptotically almost converges to TDMA as time passes instead of being much higher than TDMA in the case of  $\beta = 0$ . Meanwhile, the FU-enhanced AoI still maintains a significant performance advantage regarding regret and system usage when compared to GF and TDMA.

Fig. 6 illustrates the convergence of the estimated hyperparameter values  $q_{nk}^*$  and  $\epsilon_{b_k}^{*(n)}$ . The error is measured as the difference between the true regret of the forward algorithm using the true hyperparameter values  $q_{nk}$  and  $\epsilon$  and the regret resulting from scheduling the resources for the devices using the estimated hyperparameter values. We initialize the values of  $q_{nk}^*$  and  $\epsilon_{b_k}^{*(n)}$  and run the iterative optimization algorithm as described in section V. We solve (32) for each device using both exhaustive search and CVX, where exhaustive search results in a more accurate estimation, while CVX is much simpler and more efficient in terms of estimation time. Afterward, we run the Baum-Welsh algorithm for 40 iterations, where it converges to reasonable values for  $\epsilon_{b_k}^{*(n)}$  and  $q_{nk}^*$  that truly describe the observations. We can

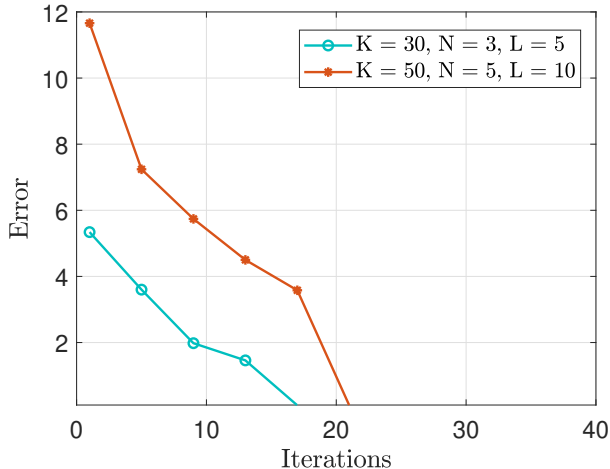
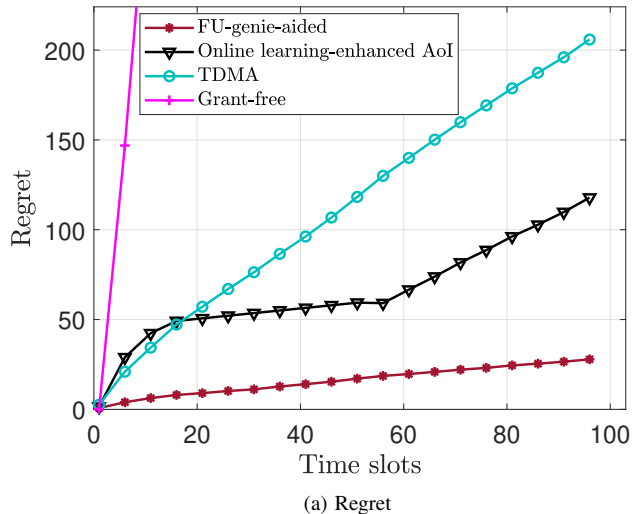


Fig. 6: The convergence of estimation of the model hyperparameters using the Baum-welsh algorithm.

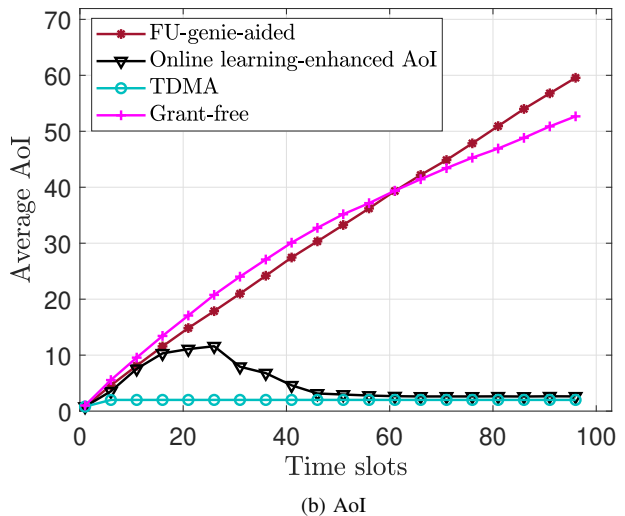
notice the convergence of the model hyperparameters after looping the algorithm for a sufficient number of iterations. Typically, the convergence is significantly faster for a small setup of the system model as the number of states and devices controls the number of the hyperparameters to be estimated. We run the mentioned estimation procedure to be used in the learning algorithm offline (FU-offline learning) and online (online learning-enhanced AoI), where the former assumes prior knowledge of enough number of observations to run the estimation upon it, whereas the latter runs the estimation algorithm online while accumulating the observations.

Fig. 7 shows the performance evaluation of the online learning-enhanced AoI algorithm in terms of regret and average AoI, respectively. As the algorithm has no prior knowledge about the states and the hyperparameters of the model, it applies the forward algorithm and the age compensation strategy based on the given set of previous observations collected at each time step. We can see in Fig. 7-(a) that the behavior of the algorithm is not efficient in the initial time steps as there are not enough observations that can describe the model and correctly estimate the model hyperparameters. Afterward, the hyperparameters estimation gets better (almost after 16 time instants) as the model collects a suitable amount of observations that truly describe the model and are used efficiently in the estimation procedure. In Fig. 7-(b), the algorithm experiences a large AoI compared to the TDMA in the initial time steps, where the age compensation strategy optimizes the age parameter  $\beta$  assuming that the prediction results are efficient enough to compensate the true high age devices. Afterward, the online learning-enhanced AoI algorithm collects enough observation to efficiently predict the model hyperparameters, where the age compensation strategy almost captures the AoI of the TDMA after 40-time instants.

Fig. 8 summarizes the regret, AoI, and system usage performance metrics when applying the proposed resource allocation schemes. It is worth mentioning that the GF results are omitted from the bar plots as it has extremely poor performance compared to all other schemes due to high number of collisions,



(a) Regret



(b) AoI

Fig. 7: Regret and AoI evaluation of the online learning-enhanced AoI algorithm.  $K = 50$  sensors,  $N = 5$  events, and  $L = 10$  available frequency resources.

and this would affect the comprehensive comparison of the schemes on the plots (namely, on the regret bar plot). It results in regret of around 3000, an AoI of 52, and system usage of 65%. The FU-feedback achieves a reduced regret to 50 times less than the GF and a slightly less system usage than the FU-genie-aided case with 2% difference. The TDMA has the best AoI results as it is considered as the fair age scheduler. Therefore, age compensation is applied within the FU-enhanced AoI algorithm that captures the AoI of the TDMA of 2.3 at the expense of slightly higher regret, where it has a 40 more regret than the FU-feedback. However, it still outperforms the regret and the system usage of TDMA and GF schemes. We can observe that the FU-baseline achieves 3 times lower regret than and 9% higher system usage than TDMA. Therefore, the FU-baseline still outperforms the TDMA and the GF resource allocation schemes regarding regret and system usage with lower computational demands.

Moreover, we fit the estimated parameters to the scheduling

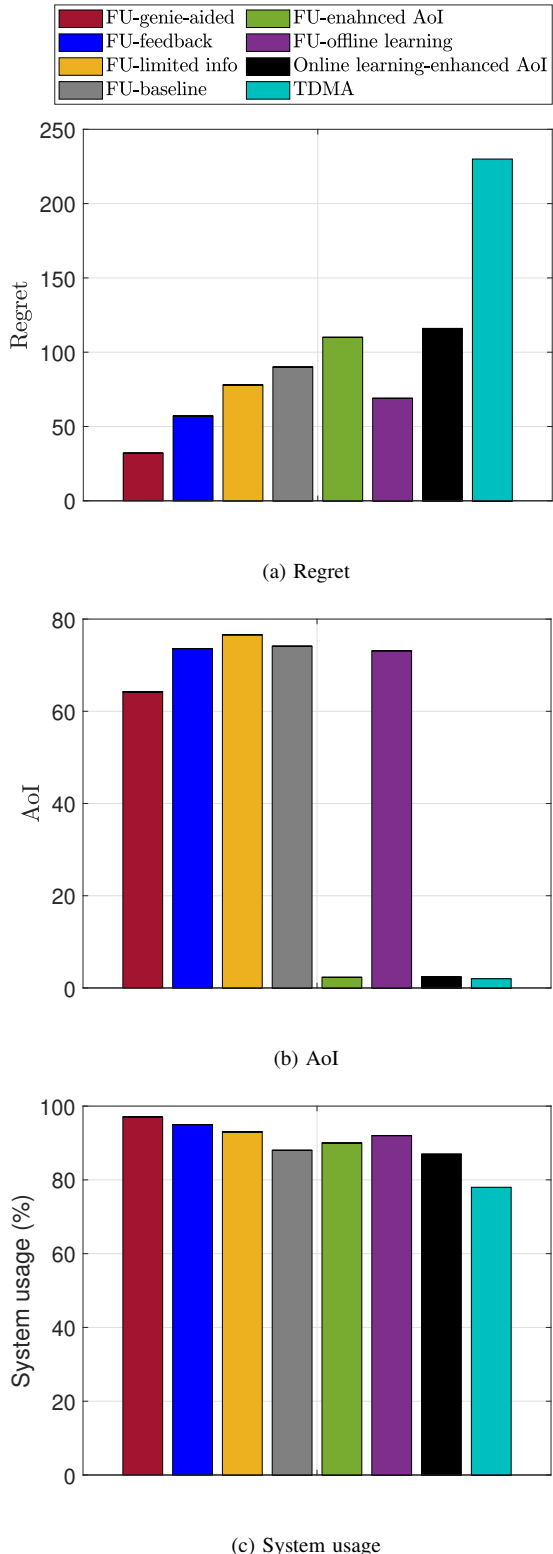


Fig. 8: Regret, AoI and system usage evaluation.  $K = 50$  sensors,  $N = 5$  events, and  $L = 10$  available frequency resources.

algorithm to calculate the model’s regret, system usage, and average AoI. We can observe that both FU-offline learning and online learning-enhanced AoI outperform the regret of the

TDMA and almost captures the regret of the FU-feedback. In addition, the online learning-enhanced AoI has almost double the regret of the FU-offline learning (65 and 120 for the FU-offline learning and the online learning-enhanced AoI, respectively) as the online version suffers from inaccurate estimation at the beginning of the simulation as there are not enough observations to be used in the estimation, whereas the offline version assumes prior knowledge of enough observations for the estimation. In addition, the online learning-enhanced AoI performs an AoI compensation step after estimating the model hyperparameters, which enables the algorithm to achieve the AoI of the TDMA while preserving the regret to still outperform the TDMA. Finally, There is an interesting analogy between the FU-limited info and the FU-offline learning results, where both algorithms suffer from missing information as the former has limited information about the actual activation of the devices and depends only on its prediction, whereas the latter relies on a collection of past observations to estimate the model hyperparameters.

## VII. CONCLUSIONS

This paper considers Markovian events which serve to model the activity of the massive deployment of IoT devices. We proposed an FU algorithm that efficiently predicts the activation pattern of the IoT devices based on the forward algorithm and grants the available resources to the devices with the highest likelihood of activation probabilities. We formulated an optimization problem that compromises a small value of the regret to minimize the AoI of the IoT devices and achieve a desirable degree of fairness. In addition, we formulated an expectation-maximization algorithm based on the Baum-Welsh procedure to estimate the system hyperparameters. Finally, we developed an online-learning version of the proposed scheme. Simulation results showed that the proposed algorithm outperforms the existing models, e.g., TDMA and GF, regarding regret, system usage efficiency, and AoI.

The proposed algorithms were much simpler than machine learning-based predictors regarding the complexity of the computations. Therefore, the proposed algorithms could be used as traffic predictors in critical applications, e.g., predictive UAV positioning [40], road safety, and other applications with low latency communications demands [6].

## REFERENCES

- [1] M. Latva-aho and K. Leppanen, “Key Drivers and Research Challenges for 6G Ubiquitous Wireless Intelligence,” *6G Flagship, University of Oulu, Finland*, Sep 2019.
- [2] N. H. Mahmood, H. Alves, O. A. López, M. Shehab, D. P. M. Osorio, and M. Latva-aho, “Six key enablers for machine type communication in 6G,” *6G Summit*, 2019.
- [3] M. Laner, P. Svoboda, N. Nikaein, and M. Rupp, “Traffic models for machine type communications,” in *ISWCS 2013; The Tenth International Symposium on Wireless Communication Systems*, 2013, pp. 1–5.
- [4] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, “Minimizing age of information in vehicular networks,” in *8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*. IEEE, 2011, pp. 350–358.
- [5] A. M. Bedewy, Y. Sun, and N. B. Shroff, “Minimizing the age of information through queues,” *IEEE Transactions on Information Theory*, vol. 65, no. 8, pp. 5215–5232, 2019.

- [6] P. Popovski, Č. Stefanović, J. J. Nielsen, E. de Carvalho, M. Angjelichinoski, K. F. Trillingsgaard, and A. Bana, "Wireless access in ultra-reliable low-latency communication (URLLC)," *IEEE Transactions on Communications*, vol. 67, no. 8, pp. 5783–5801, 2019.
- [7] A. Laya, L. Alonso, and J. Alonso-Zarate, "Is the Random Access Channel of LTE and LTE-A Suitable for M2M Communications? A Survey of Alternatives," *IEEE Communications Surveys Tutorials*, vol. 16, no. 1, pp. 4–16, First 2014.
- [8] N. H. Mahmood, R. Abreu, R. Böhneke, M. Schubert, G. Berardinelli, and T. H. Jacobsen, "Uplink grant-free access solutions for URLLC services in 5G new radio," in *2019 16th International Symposium on Wireless Communication Systems (ISWCS)*, 2019, pp. 607–612.
- [9] 3GPP, "Study on RAN improvements for machine-type communications," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 37.868, 2012.
- [10] R.-G. Cheng, J. Chen, D.-W. Chen, and C.-H. Wei, "Modeling and analysis of an extended access barring algorithm for machine-type communications in LTE-A networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 2956–2968, 2015.
- [11] S.-Y. Lien, T.-H. Liao, C.-Y. Kao, and K.-C. Chen, "Cooperative access class barring for machine-to-machine communications," *IEEE Transactions on Wireless Communications*, vol. 11, no. 1, pp. 27–32, 2012.
- [12] S. Duan, V. Shah-Mansouri, and V. W. S. Wong, "Dynamic access class barring for M2M communications in LTE networks," in *2013 IEEE Global Communications Conference (GLOBECOM)*, 2013, pp. 4747–4752.
- [13] 3GPP, "Service accessibility," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 22.011, 12 2018, version 16.4.0.
- [14] Y. Xu, G. Gui, H. Gacanin, and F. Adachi, "A survey on resource allocation for 5G heterogeneous networks: Current research, future trends, and challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 668–695, 2021.
- [15] S. Ali, N. Rajatheva, and W. Saad, "Fast uplink grant for machine type communications: Challenges and opportunities," *IEEE Communications Magazine*, vol. 57, no. 3, pp. 97–103, March 2019.
- [16] M. Shehab, A. K. Hagelskjær, A. E. Kalør, P. Popovski, and H. Alves, "Traffic prediction based fast uplink grant for massive IoT," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, 2020, pp. 1–6.
- [17] M. Laner, P. Svoboda, N. Nikaein, and M. Rupp, "Traffic Models for Machine Type Communications," in *ISWCS 2013; The Tenth International Symposium on Wireless Communication Systems*, Aug 2013, pp. 1–5.
- [18] E. Grigoreva, M. Laurer, M. Vilgelm, T. Gehrsitz, and W. Kellerer, "Coupled Markovian Arrival Process for Automotive Machine Type Communication traffic modeling," in *2017 IEEE International Conference on Communications (ICC)*, May 2017, pp. 1–6.
- [19] P. Salvo Rossi, D. Ciuonzo, and T. Ekman, "HMM-based decision fusion in wireless sensor networks with noncoherent multiple access," *IEEE Communications Letters*, vol. 19, no. 5, pp. 871–874, 2015.
- [20] A. E. Kalør, O. A. Hanna, and P. Popovski, "Random Access Schemes in Wireless Systems with Correlated User Activity," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2018, pp. 1–5.
- [21] S. M. Hasan, K. Mahata, and M. M. Hyder, "Fast uplink grant-free NOMA with sinusoidal spreading sequences," 2020.
- [22] Z. Zhou, R. Ratasuk, N. Mangalvedhe, and A. Ghosh, "Resource allocation for uplink grant-free ultra-reliable and low latency communications," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, 2018, pp. 1–5.
- [23] S. Ali, A. Ferdowsi, W. Saad, and N. Rajatheva, "Sleeping multi-armed bandits for fast uplink grant allocation in machine type communications," in *2018 IEEE Globecom Workshops (GC Wkshps)*, Dec 2018, pp. 1–6.
- [24] E. Eldeeb, M. Shehab, and H. Alves, "A learning-based fast uplink grant for massive IoT via support vector machines and long short-term memory," *IEEE Internet of Things Journal*, pp. 1–1, 2021.
- [25] O. Habachi, M.-A. Adjif, and J.-P. Cances, "Fast uplink grant for NOMA: A federated learning based approach," in *Ubiquitous Networking*, O. Habachi, V. Meghdadi, E. Sabir, and J.-P. Cances, Eds. Cham: Springer International Publishing, 2020, pp. 96–109.
- [26] I. AlQerm and B. Shihada, "A cooperative online learning scheme for resource allocation in 5G systems," in *2016 IEEE International Conference on Communications (ICC)*, 2016, pp. 1–7.
- [27] A. E. Kalør, O. Simeone, and P. Popovski, "Prediction of mmwave/thz link blockages through meta-learning and recurrent neural networks," 2021. [Online]. Available: <https://arxiv.org/abs/2106.07442>
- [28] F. Shah-Mohammadi and A. Kwasinski, "Deep reinforcement learning approach to QoE-driven resource allocation for spectrum underlay in cognitive radio networks," in *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2018, pp. 1–6.
- [29] X. Liu, J. Yu, J. Wang, and Y. Gao, "Resource allocation with edge computing in IoT networks via machine learning," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3415–3426, 2020.
- [30] D. Hejji, M. A. Talib, A. B. Nassif, Q. Nasir, and A. Bouridane, "AI-based models for resource allocation and resource demand forecasting systems in aviation: A survey and analytical study," in *2021 IEEE International Conference on Internet of Things and Intelligence Systems (IoT&IS)*, 2021, pp. 183–189.
- [31] N. Kato, B. Mao, F. Tang, Y. Kawamoto, and J. Liu, "Ten challenges in advancing machine learning technologies toward 6G," *IEEE Wireless Communications*, vol. 27, no. 3, pp. 96–103, 2020.
- [32] A. Gepperth and B. Hammer, "Incremental learning algorithms and applications," in *European Symposium on Artificial Neural Networks (ESANN)*, Bruges, Belgium, 2016. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-01418129>
- [33] C. Ren, X. Lyu, W. Ni, H. Tian, and R. P. Liu, "Distributed online learning of Fog computing under nonuniform device cardinality," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 1147–1159, 2019.
- [34] A. Kosta, N. Pappas, and V. Angelakis, "Age of information: A new concept, metric, and tool," *Foundations and Trends in Networking, Now Publishers, Inc.*, 2017.
- [35] O. Capp, E. Moulines, and T. Ryden, *Inference in Hidden Markov Models*. Springer Publishing Company, Incorporated, 2010.
- [36] S. Sefati and N. J. Navimipour, "A QoS-aware service composition mechanism in the internet of things using a hidden-markov-model-based optimization algorithm," *IEEE Internet of Things Journal*, vol. 8, no. 20, pp. 15 620–15 627, 2021.
- [37] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, Feb 1989.
- [38] S. L. C. G. Cassandras, *Introduction to Discrete Event Systems*. Springer, 2008.
- [39] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," <http://cvxr.com/cvx>, Mar. 2014.
- [40] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Mobile Unmanned Aerial Vehicles (UAVs) for Energy-Efficient Internet of Things Communications," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7574–7589, Nov 2017.