Minimizing Age of Information via Scheduling over Heterogeneous Channels

Jiayu Pan  
Department of ECE  
The Ohio State University  
Columbus, OH, USA  
pan.743@osu.edu

Ahmed M. Bedewy  
Department of ECE  
The Ohio State University  
Columbus, OH, USA  
bedewy.2@osu.edu

Yin Sun  
Department of ECE  
Auburn University  
Auburn, AL, USA  
yzs0078@auburn.edu

Ness B. Shroff  
Departments of ECE and CSE  
The Ohio State University  
Columbus, OH, USA  
shroff.11@osu.edu

ABSTRACT

In this paper, we study the problem of minimizing the age of information when a source can transmit status updates over two heterogeneous channels. Our work is motivated by recent developments in 5G mmWave technology, where transmissions may occur over an unreliable but fast (e.g., mmWave) channel or a slow reliable (e.g., sub-6GHz) channel. The unreliable channel is modeled as a time-correlated Gilbert-Elliot channel, where information can be transmitted at a high rate when the channel is in the “ON” state. The reliable channel provides a deterministic but lower data rate. The scheduling strategy determines the channel to be used for transmission with the aim to minimize the average age of information (AoI). The optimal scheduling problem is formulated as a Markov Decision Process (MDP), which in our setting poses some significant challenges because e.g., supermodularity does not hold for part of the state space. We show that there exists a multi-dimensional threshold-based scheduling policy that is optimal for minimizing the age. A low-complexity bisection algorithm is further devised to compute the optimal thresholds. Numerical simulations are provided to compare different scheduling policies.

CCS CONCEPTS

• Networks → Network performance evaluation; Network performance analysis;

KEYWORDS

Age of information; Heterogeneous channels; Scheduling policy

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MobiHoc ’21, July 26–29, 2021, Shanghai, China  
© 2021 Association for Computing Machinery.  
ACM ISBN 978-1-4503-8558-9/21/07...$15.00  
https://doi.org/10.1145/3466772.3467040

1 INTRODUCTION

The timely update of the system state is of great significance in cyber-physical systems such as vehicular networks, sensor networks, and UAV navigations. In these applications, newly generated data is more desirable than outdated data. Age of information (AoI), or simply age, was introduced as an end-to-end application-layer metric to evaluate data freshness [1, 3–6, 8, 9, 12, 14–18, 20, 21, 24, 29, 35–40, 43]. The age at time \( t \) is defined as \( \Delta(t) = t - U_t \), where \( U_t \) is the generation time of the freshest packet that has been received by time \( t \). The difference between age and classical metrics like delay and throughput is evident even in elementary queuing systems [18]. High throughput implies frequent status updates, but tends to cause a higher queuing delay that worsens timeliness. On the other hand, delay can be greatly reduced by decreasing the update frequency, which, however, may increase the age because the status is updated infrequently.

In future wireless networks, the sub-6GHz frequency spectrum is insufficient for fulfilling the high throughput demand of emerging real-time applications such as VR/AR applications, where contents must be delivered within 5–20 ms of latency, requiring a high throughput of 400–600 Mbps [30]. To address this challenge, 5G technology utilizes high-frequency millimeter wave (mmWave) bands such as 28/38 GHz, which provide a much higher data rate than sub-6GHz [31]. Recently, Verizon and Samsung demonstrated that a throughput of nearly 4Gbps was achieved in their mmWave demo system, using a 28GHz frequency band with 800MHz bandwidth [32]. However, unlike sub-6GHz spectrum bands, mmWave channels are highly unreliable due to blocking susceptibility, strong atmospheric absorption, and low penetration. Real-world smartphone experiments have shown that even obstructions by hands could significantly degrade the mmWave throughput [22]. One solution to mitigate this effect is to let sub-6GHz coexist with mmWave to form two heterogeneous channels, so that the user equipment can offload data to sub-6GHz when mmWave communications are unfeasible [2, 26, 27, 33]. Some work has already been done based
on mmWave/sub-6GHz heterogeneous networks [10, 13]. However, 
how to improve information freshness in such hybrid networks has 
remained largely unexplored.

In this study, we consider a hybrid status updating system where 
a source can transmit the update packets over either an unreliable 
but fast mmWave channel or a slow reliable sub-6GHz channel. Our 
objective is to find a dynamic channel scheduling policy that mini-
mizes the long-term average expected age. The main contributions 
of this paper are stated as follows:

- The optimal scheduling problem for minimizing the age over 
heterogeneous channels is formulated as a Markov Decision 
Process (MDP). The state transition of this MDP is compli-
cated for two reasons: (i) the two channels have different 
data rates and packet transmission times, and (ii) the state of 
the unreliable mmWave channel is correlated over time. We 
prove that there exists a multi-dimensional threshold-based 
scheduling policy that is optimal. This optimality result holds 
for all possible values of the channel parameters. Supermod-
ularity [41] has been one of the tools used to prove this result. 
Because of the complicated state transitions, the super-
modular property only holds in a part of the state space, 
which is a key difference from the scheduling problems con-
sidered earlier in, e.g., [1, 19, 23, 28, 35, 38, 42]. We have 
developed additional techniques to show that the threshold-
base scheduling policy is optimal in the rest part of the 
state space where supermodularity does not hold.

- Further, we show that the thresholds of the optimal schedul-
ing policy can be evaluated efficiently, by using closed-form expressions or a low-complexity bisection search algorithm. 
Compared with the algorithms for calculating the thresholds and optimal scheduling policies in, e.g., [1, 19, 23, 28, 35, 38, 42], our solution algorithms have much lower computational complexities.

- In the special case that the state of the unreliable mmWave 
channel is independent and identically distributed (i.i.d.) over time, the optimal scheduling policy is shown to possess an 
even simpler form.

- Finally, numerical results show that the optimal policy can 
reduce the age compared with several other policies.

2 RELATED WORKS

Age of information has become a popular research topic in recent 
years, e.g., [1, 3–6, 8, 9, 12, 14–18, 20, 21, 24, 29, 35–40, 43]. A com-
prehensive survey of the area was recently provided in [43]. First, 
there has been substantial work on age performance analysis in 
queueing disciplines [4, 5, 8, 9, 15, 18]. Average age and peak age in 
elementary queueing systems were analyzed in [9, 15, 18]. A similar 
setting with Gilbert-Elliot sampler or Gilbert-Elliot server was con-
sidered in [8]. A Last-Generated, First-Served (LGFS) policy was 
shown (near) optimal in general single source, multiple servers, and 
multihop networks with arbitrary generation and arbitrary arrival 
process [4, 5]. These results were extended to the multi-source 
multi-server regime in [37]. Next, there has been a significant effort 
in age-optimal sampling [3, 24, 35, 36, 38]. The optimal sampling 
policy was provided for minimizing a monotonic age function in 
[24, 35, 38]. Sampling and scheduling in multi-source systems were 
analyzed where the optimal joint problem could be decoupled into 
maximum age first (MAF) scheduling [37] and an optimal sam-
ppling problem in [3]. Finally, age in wireless networks has been 
substantially explored in [14, 16, 17, 20, 21, 29, 40]. Scheduling in a 
broadcast network with random arrival was provided where whittle 
index policy can achieve (near) age optimality [14]. Some other age-
optimal scheduling works for cellular networks were considered in 
[16, 17, 21, 39, 40]. A class of age-optimal scheduling policies was 
analyzed in the asymptotic regime when the number of sources and 
channels both grow to infinity [29]. An age minimization multi-path 
routing strategy was introduced in [20].

However, the age-optimal problem via heterogeneous channels 
has been largely unexplored yet. To the best of our knowledge, tech-
nical models similar to ours were reported in [1, 12]. Their study 
assumed that the first channel is unreliable but consumes a lower 
cost, and the second channel has a steady connection with the same 
delay but consumes a higher cost. They derived the scheduling policy 
for the trade-off between age performance and cost. Our 
study is significantly different from theirs in two aspects: (i) The 
study in [1, 12] shows the optimality of the threshold type policy 
efficiently solves the optimal threshold when the first channel is 
i.i.d. [1], but our work allows a Markovian channel which is a 
geeneralization of the i.i.d. case. (ii) In addition to allowing mmWave 
to be unreliable, our study assumes that sub-6GHz has a larger delay 
than mmWave since this assumption complies with the property of 
dual mmWave/sub-6GHz channels in real applications. These two 
differences between mmWave and sub-6GHz make the MDP for-
mulation more complex. Thus, most of the well-known techniques 
that show a nice structure of the optimal policy or even solve the 
opimal policy with low complexity (e.g., [1, 19, 23, 28, 35, 38, 42]) 
do not apply to our model.

3 SYSTEM MODEL AND PROBLEM FORMULATION

3.1 System Models

Consider a single-hop network as illustrated in Fig. 1, where a 
source sends status update packets to the destination. We assume 
that time is slotted with slot index \( t \in \{0, 1, 2, ..., \} \). The source 
can generate a fresh status update packet at the beginning of each 
time slot. The packets can be transmitted either over the mmWave 
channel or over the sub-6GHz channel. The packet transmission 
time of the mmWave channel is 1 time slot, whereas the packet 
transmission time of the sub-6GHz channel is \( d \) time slots (\( d \geq 2 \)) 
because of its lower data rate.

The mmWave channel, called Channel 1, follows a two-state 
Gilbert-Elliot model that is shown in Fig. 2. We say that Channel 1 
is ON in time slot \( t \), denoted by \( I_1(t) = 1 \), if the packet is successfully 
transmitted to the destination in time slot \( t \); otherwise Channel 1 is 
said to be OFF, denoted by \( I_1(t) = 0 \). If a packet is not successfully 
transmitted, then it is dropped, and a new status update packet is 
generated at the beginning of the next time slot. The self transition 
probability of the ON state is \( p \), and the self transition probability of 
the OFF state is \( q \), where \( 0 < q < 1 \) and \( 0 < p < 1 \). We assume 
that the source has access to the state of Channel 1, but with one

\[^{1}\text{If } d = 1, \text{ one can readily see that it is better to choose sub-6GHz than mmWave. Thus, in this paper we study the nontrivial case of } d \geq 2.\]
As mentioned above, the packet transmission time of Channel 2 is 1) or sub-6GHz (Channel 2) for transmission over time. Following the application settings in [2, 10, 26, 27, 33], a packet can be transmitted using only one channel at a time, i.e., the two channels cannot be used simultaneously. The scheduler decides which channel to be used to transmit a packet at each time slot. We also assume that the scheduler can choose idle (neither channel) since it has been shown that channel idling could reduce the average age in the system [3, 35, 38]. The scheduling decision at the beginning of time slot \( t \) is denoted by \( u(t) \in \{1, 2, \text{none}\} \). The action \( u(t) = 1 \) or 2 means that the source generates a packet and assigns it to Channel 1 or Channel 2, respectively. The action \( u(t) = \text{none} \) means that no new packet is assigned to any channel at time slot \( t \). Hence, \( u(t) = \text{none} \) can occur if (i) a packet is assigned to Channel 2 earlier and has not completed its transmission, i.e., \( l_2(t) \in \{1, 2, \ldots, d - 1\} \) such that no packet can be assigned for transmission, or (ii) \( l_2(t) = 0 \), but both channels are kept idle on purpose.

The age of information (AoI) \( \Delta(t) \) is the time difference between the current time slot \( t \) and the generation time of the freshest delivered packet [18]. By this definition, when a packet is delivered, the age drops to the transmission time of the delivered packet. Specifically, if Channel 1 is selected in time slot \( t \) and Channel 1 is ON, then the age drops to 1 at time slot \( t + 1 \). If the remaining service time of Channel 2 at time slot \( t \) is 1, then age drops to \( d \) at time slot \( t + 1 \). When there is no packet delivery at time slot \( t \), the age increases by one time slot. Hence, the time-evolution of the age is given by

\[
\Delta(t + 1) = \begin{cases} 
1 & \text{if } u(t) = 1 \text{ and } l_1(t) = 1, \\
\Delta(t) + 1 & \text{otherwise.}
\end{cases}
\]

### 3.2 Problem Formulations

We use \( \pi = \{u(0), u(1), \ldots\} \) to denote a scheduling policy. A scheduling policy is said to be **admissible** if (i) \( u(t) = \text{none} \) whenever \( l_2(t) \geq 1 \) and (ii) \( u(t) \) is determined by the current and history information that is available at the scheduler. Let \( \Delta_\pi(t) \) denote the AoI induced by policy \( \pi \). The expected time-average age of policy \( \pi \) is

\[
\limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E[\Delta_\pi(t)].
\]

Our objective in this paper is to solve the following optimal scheduling problem for minimizing the expected time-average age:

\[
\Delta_{\text{opt}} = \inf_{\pi \in \Pi} \limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E[\Delta_\pi(t)],
\]

where \( \Pi \) is the set of all admissible policies. Problem (2) can be equivalently expressed as an average-cost MDP problem [7, 28], which is illustrated below:

- **Markov State**: The system state in time slot \( t \) is defined as

\[
s(t) = (\Delta(t), l_1(t-1), l_2(t)),
\]

![Figure 1: The system model for status updates in heterogeneous channels. The scheduler chooses mmWave (Channel 1) or sub-6GHz (Channel 2) for transmission over time.](image1)

![Figure 2: The Gilbert-Elliot model for Channel 1.](image2)

<table>
<thead>
<tr>
<th>Table 1: Value of State Transition Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_{s\to s'}(u) )</td>
</tr>
<tr>
<td>( p )</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>( 1 - p )</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>( q )</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>( 1 - q )</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>0</td>
</tr>
</tbody>
</table>
As mentioned in Section 3.2, the action space of the MDP allows the minimum non-zero waiting time is one time slot which is the same as the zero wait policy (transmitting immediately after the previous transmission). Let $S$ denote the state space which is countably infinite. The time-evolution of $\Delta(t)$ is determined by the state and action in time slot $t$.

- **Action**: As mentioned before, if Channel 2 is busy (i.e., $l_2(t) > 0$), the scheduler always chooses an idle action, i.e., $u(t) = \text{none}$. Otherwise, the action $u(t) \in \{1, 2, \text{none}\}$.
- **Cost function**: Suppose that a decision $u(t)$ is applied at a time slot $t$, we encounter a cost $C(s(t), u(t)) = \Delta(t)$.
- **Transition probability**: We use $P_{sw}(u)$ to denote the transition probability from state $s$ to $s'$ for action $u$. The value of $P_{sw}(u)$ is summarized in Table 1. See our technical report [25] for the explanation of Table 1.

4 MAIN RESULTS

In this section, we show that there exists a threshold-type policy that solves Problem (2). We then provide a low-complexity algorithm to obtain the optimal policy and optimal average age.

4.1 Optimality of threshold-type policies

As mentioned in Section 3.2, the action space of the MDP allows $u(t) = \text{none}$ even if Channel 2 is idle, i.e., $l_2(t) = 0$. In the following lemma, we show that the action $u(t) = \text{none}$ can be abandoned when $l_2(t) = 0$. Define

$$
\Pi' = \{\pi \in \Pi : u(t) \neq \text{none}, \text{if } l_2(t) = 0\}. \tag{4}
$$

**Lemma 1.** For any $\pi \in \Pi$, there exists a policy $\pi' \in \Pi'$ that is no worse than $\pi$.

**Proof.** See our technical report [25].

By Lemma 1, the scheduler only needs to choose from the actions $u(t) = 1$ or 2 when $l_2(t) = 0$. This lemma simplifies the MDP problem.

**Remark 1.** In [3, 35, 38], the authors showed that in certain scenarios, the zero wait policy (transmitting immediately after the previous update has been received) might not be optimal. However, in our model, the zero wait policy is indeed optimal. The reason is that in our model, the minimum non-zero waiting time is one time slot which is the same as the delay of Channel 1. If $l_2(t) = 0$, it is better to choose Channel 1 than keeping both channels idle, because, by choosing Channel 1, fresh packets could be delivered over Channel 1.

For the ease of description, we divide the possible values of channel parameters $(p, q, d)$ into four complementary regions $B_1, \ldots, B_4$.

**Definition 1.** The regions $B_1, \ldots, B_4$ are defined as

$$
\begin{align*}
B_1 &= \{(p, q, d) : F(p, q, d) \leq 0, H(p, q, d) \leq 0\}, \\
B_2 &= \{(p, q, d) : F(p, q, d) > 0, G(p, q, d) \leq 0\}, \\
B_3 &= \{(p, q, d) : F(p, q, d) > 0, G(p, q, d) > 0\}, \\
B_4 &= \{(p, q, d) : F(p, q, d) \leq 0, H(p, q, d) > 0\},
\end{align*}
$$

where $\Delta(t) \in \{1, 2, 3, \ldots\}$ is the AoI in time slot $t$, $l_1(t) \in \{0, 1\}$ is the ON-OFF state of Channel 1 in time slot $t$, and $l_2(t) \in \{0, 1, \ldots, d-1\}$ is the remaining transmission time of Channel 2 at the beginning of time slot $t$. Let $\delta$ denote the state space which is countably infinite. The time-evolution of $\delta(t)$ is determined by the state and action in time slot $t$.

In [3, 35, 38], the authors showed that in certain scenarios, the zero wait policy (transmitting immediately after the previous update has been received) might not be optimal. However, in our model, the zero wait policy is indeed optimal. The reason is that in our model, the minimum non-zero waiting time is one time slot which is the same as the delay of Channel 1. If $l_2(t) = 0$, it is better to choose Channel 1 than keeping both channels idle, because, by choosing Channel 1, fresh packets could be delivered over Channel 1.

For the ease of description, we divide the possible values of channel parameters $(p, q, d)$ into four complementary regions $B_1, \ldots, B_4$.

**Definition 1.** The regions $B_1, \ldots, B_4$ are defined as

$$
\begin{align*}
B_1 &= \{(p, q, d) : F(p, q, d) \leq 0, H(p, q, d) \leq 0\}, \\
B_2 &= \{(p, q, d) : F(p, q, d) > 0, G(p, q, d) \leq 0\}, \\
B_3 &= \{(p, q, d) : F(p, q, d) > 0, G(p, q, d) > 0\}, \\
B_4 &= \{(p, q, d) : F(p, q, d) \leq 0, H(p, q, d) > 0\},
\end{align*}
$$

where the values of $\delta$, $\delta(t)$, $\lambda$, and $\lambda(t)$ are defined as

$$
\begin{align*}
\delta &= \frac{1}{\lambda} - d, \\
\delta(t) &= \frac{t}{\lambda} - d, \\
\lambda &= \frac{1}{1 - \beta}, \\
\lambda(t) &= \frac{t}{1 - \beta}.
\end{align*}
$$

In this study, the regions $B_1 - B_4$ will serve as the sufficient conditions for optimality of threshold type policy. Note that the inequality $1/(1 - \beta) > d$ also represents a comparison between the channel delay $d$ and the average length of OFF period given that the last channel state is OFF. Similarly, $1 - \frac{1}{(1 - \beta) + 1 - d}$ represents a comparison between $d$ and the average length of OFF period given that the last channel state is ON. Finally, $(1 - q)/(1 - p) + 1 - d > 0$ represents a comparison between $d$ and the average delay of Channel 1. These comparisons are the interpretations of all the boundary functions $F, G, H$ of the regions $B_1 - B_4$. The four regions $B_1, \ldots, B_4$ are depicted in Fig. 3.

Consider a stationary policy $\mu(\delta, l_1, l_2)$. As mentioned in Lemma 1, when $l_2 = 0$, the decision $\mu(\delta, l_1, 0)$ can be 1 (Channel 1) or 2 (Channel 2). Given the value of $l_1$, $\mu(\delta, l_1, 0)$ is said to be non-decreasing in the age $\delta$, if

$$
\mu(\delta, l_1, 0) = \begin{cases} 
1 & \text{if } \delta < \lambda; \\
2 & \text{if } \delta \geq \lambda.
\end{cases}
$$

Conversely, $\mu(\delta, l_1, 0)$ is said to be non-increasing in the age $\delta$, if

$$
\mu(\delta, l_1, 0) = \begin{cases} 
2 & \text{if } \delta < \lambda; \\
1 & \text{if } \delta \geq \lambda.
\end{cases}
$$

One can observe that scheduling policies in the form of (7) and (8) are both with a threshold-type, where $\lambda$ is the threshold on the age $\delta$ at which the value of $\mu(\delta, l_1, 0)$ changes.
One possible solution to Problem (2) is of a special threshold-type structure, as stated in the following theorem:

**Theorem 1.** There exists an optimal solution \( \mu^*(\delta, l_1, 0) \) to Problem (2), which satisfies the following properties:

(a) if \((p, q, d) \in B_1\), then \( \mu^*(\delta, 0, 0) \) is non-increasing in the age \( \delta \) and \( \mu^*(\delta, 1, 0) \) is non-increasing in the age \( \delta \);

(b) if \((p, q, d) \in B_2\), then \( \mu^*(\delta, 0, 0) \) is non-decreasing in the age \( \delta \) and \( \mu^*(\delta, 1, 0) \) is non-increasing in the age \( \delta \);

(c) if \((p, q, d) \in B_3\), then \( \mu^*(\delta, 0, 0) \) is non-decreasing in the age \( \delta \) and \( \mu^*(\delta, 1, 0) \) is non-increasing in the age \( \delta \);

(d) if \((p, q, d) \in B_4\), then \( \mu^*(\delta, 0, 0) \) is non-increasing in the age \( \delta \) and \( \mu^*(\delta, 1, 0) \) is non-decreasing in the age \( \delta \).

**Proof.** Please see Section 7.2 for the proof.

As is shown in Theorem 1, for all possible parameters \( p, q, d \) of the two channels, the optimal action \( \mu^*(\delta, l_1, 0) \) of channel selection is a monotonic function of the age \( \delta \). Whether \( \mu^*(\delta, l_1, 0) \) is non-decreasing or non-increasing in \( \delta \) depends on the channel parameters \( (p, q, d) \) and the previous state \( l_1 \) of Channel 1.

The study in [1] assumed that the first channel is unreliable and consumes a lower cost, and the second channel has a steady connection with the same delay but a higher cost. They studied the scheduling policy for the trade-off between age performance and cost. The optimal scheduling policy in Theorem 1 is quite different from that in [1]: The study in [1] assumes the first channel to be i.i.d., but our result allows a Markovian Channel 1, which is a generalization of the i.i.d. case. The optimal scheduling policy in [1] is non-decreasing in age since the first channel is inferior to the second channel. However, the optimal policy in our study can be non-decreasing or non-increasing since the two channels (i.e., Channel 1 and 2) have their own advantages. In conclusion, our study allows for general channel parameters and applies to all types of comparisons between Channel 1 and Channel 2, and our policy can be non-increasing in some regions and non-decreasing in other regions.

### 4.2 Insights Behind the Regions \( B_1 - B_4 \)

The regions \( B_1 - B_4 \) were introduced in Theorem 1 for proving that the action value function \( Q(s, u) \) is *supermodular* or *submodular*, where \( s = (\delta, l_1, 0) \) denotes the state of the MDP and \( u \) is the action. For example, in the case of \( l_1 = 0 \), if \( 1/(1-p) > d \) and \( 1/q \leq d \) (i.e., \((p, q, d) \in B_2\)), Lemma 8 in our technical report showed that \( Q(\delta, 0, u) \) is submodular in \((\delta, u)\) (in the discounted case). As a result, the optimal action \( \mu^*(\delta, 0, u) \) is increasing in \( \delta \).

However, in the case \( l_1 = 1 \) of Theorem 1, there are additional technical challenges: For example, if \((p, q, d) \in B_2\), we were unable to use \( 1/(1-p) > d \) and \( 1/q \leq d \) (where are provided in the definition of \( B_2 \)) to prove that \( Q(\delta, 1, 0, u) \) is super-modular or submodular. A new method was developed in Lemma 9 in our Appendices to conquer this challenge. Technically, super-/sub-modularity is a sufficient but not necessary condition for the monotonicity of \( \mu^*(\delta, l_1, 0) \). In some part of the state space \((l_1 = 0)\), we proved super-/sub-modularity. In the rest part of the state space \((l_1 = 1)\), neither super-modularity nor sub-modularity may hold, but we were able to show that the optimal decision \( \mu^*(\delta, 1, 0) \) does not change with \( \delta \) when \( l_1 = 1 \). By this, we proved the monotonicity of \( \mu^*(\delta, 1, 0) \) for all cases, without requiring \( Q(s, u) \) to be supermodular or submodular over the entire state space.

Thus, one technical contribution of the paper is: we proved that the optimal action \( \mu^*(\delta, 1, 0) \) is monotonic in \( \delta \) even if super-/submodularity does not hold. This is a key difference from prior studies, e.g., [1, 19, 23, 28, 35, 38], where super-modularity (or sub-modularity) holds for the entire state space.

### 4.3 Optimal Scheduling Policy

According to Theorem 1, \( \mu^*(\delta, 0, 0) \) and \( \mu^*(\delta, 1, 0) \) are both threshold-type, so there are two thresholds. We use \( \lambda^*_0 \) and \( \lambda^*_1 \) to denote the thresholds of \( \mu^*(\delta, 0, 0) \) and \( \mu^*(\delta, 1, 0) \), respectively.

**Theorem 2.** An optimal solution to (2) is presented below for the 2 regions \( B_1, B_2 \) of the channel parameters:

(a) If \((p, q, d) \in B_1\), then the optimal scheduling policy is

\[
\mu^*(\delta, 0, 0) = \begin{cases} 
1 & \text{if } \delta < \lambda^*_0 \\
2 & \text{if } \delta \geq \lambda^*_0
\end{cases}
\]

\[
\mu^*(\delta, 1, 0) = \begin{cases} 
2 & \text{if } \delta < \lambda^*_1 \\
1 & \text{if } \delta \geq \lambda^*_1
\end{cases}
\]

(b) If \((p, q, d) \in B_2\), then the optimal scheduling policy is

\[
\mu^*(\delta, 0, 0) = \begin{cases} 
1 & \text{if } \delta < \lambda^*_0 \\
2 & \text{if } \delta \geq \lambda^*_0
\end{cases}
\]

\[
\mu^*(\delta, 1, 0) = \begin{cases} 
2 & \text{if } \delta < \lambda^*_1 \\
1 & \text{if } \delta \geq \lambda^*_1
\end{cases}
\]

where \( \lambda^*_0 \) is unique, but \( \lambda^*_1 \) may take multiple values, given by

\[
\begin{align*}
\lambda^*_0 &= s_1(\beta_1), & \lambda^*_1 &= 1 & \text{if } \bar{\Lambda}_{opt} = \beta_1, \\
\lambda^*_0 &= s_2(\beta_2), & \lambda^*_1 &= 1 & \text{if } \bar{\Lambda}_{opt} = \beta_2, \\
\lambda^*_0 &= 1, & \lambda^*_1 &\in \{2, 3, \ldots, d\} & \text{if } \bar{\Lambda}_{opt} = f_0/g_0, \\
\lambda^*_0 &= 1, & \lambda^*_1 &\in \{1, 2, \ldots\} & \text{if } \bar{\Lambda}_{opt} = (3/2)d - 1/2,
\end{align*}
\]

\[
\Lambda_{opt} = \min \left\{ \beta_1, \beta_2, \frac{f_0}{g_0}, \frac{3}{2}d - \frac{1}{2} \right\}
\]

\( s_1(\cdot), s_2(\cdot), \beta_1, \) and \( \beta_2 \) are given in Definition 2 below, and

\[
f_0 = q \sum_{i=1}^{d} i + (1-q) \sum_{i=1}^{d} i + \left( \frac{q'q + b}{1-b} + 1 \right) \sum_{i=1}^{d} i,
\]

\[
g_0 = \frac{q'b + b}{1-b}d + d + 1,
\]

\[
\begin{bmatrix} b' \\ b \end{bmatrix} = \begin{bmatrix} q & 1-q & d & 0 \\ 1-p & p & 1 & 0 \end{bmatrix}
\]

**Proof.** Please see Section 7.3 for the details.

The results for the regions \( B_1 \) and \( B_4 \) are of similar forms and are relegated to our technical report [25]. The result of Theorem 2(a) is simple: if \((p, q, d) \in B_1\), the optimal policy always chooses Channel 1. However, Theorem 2(b) contains a number of cases. For each case, the optimal thresholds \( \lambda^*_0 \) and \( \lambda^*_1 \) can be either expressed in closed-form, or computed by using a low-complexity bisection search method to compute the root of (19) given in below.
The values \( \beta_1, \beta_2 \) used in Theorem 2(b) are the root of
\[
 f_i(s_i(\beta_i)) - \beta_i g_i(s_i(\beta_i)) = 0, \quad i \in \{1, 2\},
\]
where
\[
 s_1(\beta_i) = \max \left\{ \left\lceil \frac{-k_1(\beta_i)}{1 - d(1 - p)} \right\rceil, d \right\}, \quad (20)
\]
\[
 s_2(\beta_2) = \max \left\{ \min \left\{ \left\lceil \frac{-k_2(\beta_2)}{1 - d(1 - p)} \right\rceil, d \right\}, 1 \right\}, \quad (21)
\]
\[
 k_i(\beta_i) = \ell_i - \beta_i \omega_i, \quad i \in \{1, 2\},
\]
and \([x]\) is the smallest integer that is greater or equal to \( x \). For the ease of presentation, \& closed-form expressions of \( f_i(\cdot) \), \( g_i(\cdot) \), \( \ell_i \), and \( \omega_i \) for \( i = 1, 2 \) are relegated to our technical report [25].

For notational simplicity, we define
\[
 h_i(\beta) = f_i(s_i(\beta)) - \beta_i g_i(s_i(\beta)), \quad i \in \{1, 2\},
\]
(23)
The functions \( h_1(\beta), h_2(\beta) \) have the following nice property:

**Lemma 2.** For all \( i \in \{1, 2\} \), the function \( h_i(\beta) \) satisfies the following properties:

1. \( h_i(\beta) \) is continuous, concave, and strictly decreasing on \( \beta \);
2. \( h_i(0) > 0 \) and \( \lim_{\beta \to \infty} h_i(\beta) = -\infty \).

**Proof.** See our technical report [25].

Lemma 2 implies that (19) has a unique root on \([0, \infty)\). Therefore, we can use a low-complexity bisection method to compute \( \beta_1, \beta_2 \), as illustrated in Algorithm 1.

Note that the structure of Lemma 2 is motivated by Lemma 2 in [24] and Lemma 2 in [36]. In [24] and [36], since the channel is error free, the age state at the end of each transmission is independent with history information. Thus, Lemma 2 in [24] and Lemma 2 in [36] are related with a per-sample (single transmission) control. However, our study does not have such a property and thus Lemma 2 arises from solving the long term average cost of the threshold type policy.

The advantage of Theorem 2 is that the solution is easy to implement. In Theorem 2(a), we showed that the optimal policy is a constant policy that always chooses Channel 1. In Theorem 2(b), \( \bar{\Lambda}_{\text{opt}} \) is expressed as the minimization of only a few precomputed values, and the optimal policy (or the thresholds) are then provided based on which value that \( \bar{\Lambda}_{\text{opt}} \) is equal to.

Since we can use a low complexity algorithm such as bisection method to obtain \( \beta_1, \beta_2 \) in Theorem 2(b), Theorem 2 provides a solution that has much lower complexity than the state-of-the-art solution such as value iteration or Monte Carlo simulation.

**4.4 Insights behind Theorem 2**

In Theorem 1, we have successfully characterized the threshold structure for an optimal policy in region \( B_1, \ldots, B_4 \). A threshold type policy is fully identified by its thresholds \( \lambda_0, \lambda_1 \), where \( \lambda_0 \) is the threshold given that previous state of Channel 1 is \( \text{OFF} \) (i.e., \( l_1 = 0 \)) and \( \lambda_1 \) is the threshold given that previous state of Channel 1 is \( \text{ON} \) (i.e., \( l_1 = 1 \)). Thus, for a given region \( B_i \) (\( i = 1, \ldots, 4 \)), the MDP problem (2) reduces to

\[
 \bar{\Lambda}_{\text{opt}} = \min_{\lambda_0 \in \mathbb{R}, \lambda_1 \in \mathbb{R}^+} \bar{\Lambda}_i(\lambda_0, \lambda_1), \quad (24)
\]

where \( \bar{\Lambda}_i(\lambda_0, \lambda_1) \) is the long term average cost of the threshold type policy such that: (1) the threshold (monotone) structure is determined by Theorem 1 and \( B_i \); (2) the thresholds are \( \lambda_0, \lambda_1 \). Note that a threshold type policy is stationary and thus can be modeled as a discrete time Markov chain (DTMC). Then, \( \bar{\Lambda}_i(\lambda_0, \lambda_1) \) can be solved by deriving the stationary distribution of the DTMC.

From Lemma 10 and Lemma 11 in Section 7.3, if \( (p, q, d) \in B_1 \), then \( \mu^*(1, 0, 0) = 1 \) and \( \mu^*(1, 1, 0) = 1 \). According to Theorem 1(a), if \( (p, q, d) \in B_1 \), then the optimal policy \( \mu^*(\delta, 0, 0) \) is non-increasing and \( \mu^*(\delta, 1, 0) \) is non-increasing. Thus, \( \mu^*(\delta, 0, 0) = 1 \) and \( \mu^*(\delta, 1, 0) = 1 \) for all \( \delta \). That is, the optimal policy is always choosing Channel 1. Since the DTMC for always choosing Channel 1 is easy to analyze, we omit the derivation steps and get

\[
 \bar{\Lambda}_{\text{opt}} = \bar{\Lambda}_1(1, 1) = \frac{(1 - q)(2 - p) + (1 - p)^2}{(2 - q - p)(1 - p)}, \quad (25)
\]

This result directly implies Theorem 2(a).

While the result of case \( (p, q, d) \in B_1 \) seems easy to describe, the result of the case \( (p, q, d) \in B_2 \) is not, because the optimal policy is no longer constant in \( \delta \). We now provide the sketch of the proof idea when \( (p, q, d) \in B_2 \).

First, by deriving the stationary distributions of some DTMCs with different thresholds, we have found that

\[
 \bar{\Lambda}_2(\lambda_0, 1) = \begin{cases} f_1(\lambda_0)/g_1(\lambda_0) & \lambda_0 \in \{d + 1, \ldots\}, \\ f_2(\lambda_0)/g_2(\lambda_0) & \lambda_0 \in \{2, \ldots, d\}. \end{cases} \quad (26)
\]

\[
 \bar{\Lambda}_2(1, \lambda_1) = \begin{cases} (3/2)d - 1/2 & \lambda_1 \in \{d + 1, \ldots\}, \\ 3/\lambda_1 & \lambda_1 \in \{1, \ldots, d\}. \end{cases} \quad (27)
\]

Note that (24) is a two-dimensional optimization problem in \((\lambda_0, \lambda_1)\). However, (24) can be reduced to a couple of one-dimensional optimization problem in \( \lambda_0 \). The reason is that the threshold type policies with different \( \lambda_1 \) may have the same DTMC; see (27) for an example.

Then, the optimal average age \( \bar{\Lambda}_{\text{opt}} \) satisfies

\[
 \bar{\Lambda}_{\text{opt}} = \min \left\{ \bar{\beta}_1, \bar{\beta}_2; \frac{f_1(\lambda_0)}{g_1(\lambda_0)}, \frac{f_2(\lambda_0)}{g_2(\lambda_0)} \right\}. \quad (28)
\]

where \( \bar{\beta}_1, \bar{\beta}_2 \) are defined as the solution to two one-dimensional problems:

\[
 \bar{\beta}_1 = \min_{\lambda_0 \in \{d + 1, \ldots\}} \frac{f_1(\lambda_0)}{g_1(\lambda_0)}, \quad (29)
\]

\[
 \bar{\beta}_2 = \min_{\lambda_0 \in \{1, \ldots, d\}} \frac{f_2(\lambda_0)}{g_2(\lambda_0)}. \quad (30)
\]

Note that (26), (27) do not cover all of the values set of \( \lambda_1, \lambda_0 \). However, only considering the 4 types of DTMC described in (26), (27) is sufficient to solve (24) for \( i = 2 \). The proof of this statement is relegated to our technical report [25].
Since \( f_1(\lambda_0) \) and \( g_1(\lambda_0) \) have a complicated structure, optimizing \( f_1(\lambda_0)/g_1(\lambda_0) \) in (29) and (30) is challenging. However, using Dinkelbach’s method [11], we can change the problem into a two-layer problem. One of our technical contributions is that the inner-layer problem is shown to be unimodal, thus we can derive an exact solution. Thus, we only need a bisection algorithm for the outer-layer, i.e., solving the roots of the equations \( h_1(\beta) = 0, h_2(\beta) = 0 \) in (19). We have

\[
\beta_i^* = \beta_i, \ i \in \{1, 2\}.
\]

The studies in [24, 35, 36] also derive an exact solution to their stopping rules [24, 36] or stochastic convex optimization [35], which is different with our study. Thus, Theorem 2(b) is solved by (26) – (31).

### 4.5 Optimal Scheduling policy for i.i.d. Channel

We finally consider a special case in which Channel 1 is i.i.d., i.e., \( p + q = 1 \). In i.i.d. case, Theorem 2 reduces to the following:

**Corollary 1.** Suppose \( p + q = 1 \), i.e., Channel 1 is i.i.d., then
(a) If \( 1 - p \geq 1/d \), then the optimal policy is always choosing Channel 1. In this case, the optimal objective value of (2) is \( \Delta_{\text{opt}} = 1/(1 - p) \).
(b) If \( 1 - p < 1/d \), then the optimal policy is non-decreasing in age and the optimal thresholds \( \lambda_0^* = \lambda_1^* \). The threshold \( \lambda_0^* \) may take multiple values, given by

\[
\begin{align*}
\lambda_0^* &\in \{1, 2, \ldots, d\} \quad \text{if } \Delta_{\text{opt}} = (3/2)d - 1/2, \\
\lambda_0^* &= s_1(\beta_1) \quad \text{if } \Delta_{\text{opt}} = \beta_1,
\end{align*}
\]

\( \Delta_{\text{opt}} \) is the optimal objective value of (2), determined by

\[
\Delta_{\text{opt}} = \min \{ \beta_1, \frac{3}{2}d - \frac{1}{2} \}.
\]

**Proof.** See our technical report [25].

If Channel 1 is i.i.d., then the state information of Channel 1 in the previous time slot should not affect the scheduling decision in the current time slot. Thus, we have only one threshold, i.e., \( \lambda_0^* = \lambda_1^* \).

Corollary 1(a) suggests that if the transmission rate of Channel 1 is larger than the rate of Channel 2 (which is \( 1/d \)), then the age-optimal policy always chooses Channel 1. Corollary 1(b) implies that if the transmission rate of Channel 1 is smaller than the rate of Channel 2, then the age-optimal policy is non-decreasing threshold-type on age.
5 NUMERICAL RESULTS

According to Corollary 1, $\lambda^*_d$ is the optimal threshold in i.i.d. channel. We provide $\lambda^*_d$ with the change of $p$ for $d = 10, 20, 50$ respectively. From Fig. 4, the optimal threshold diverges to boundary $p^* = 0.9, 0.95, 0.98$ respectively. As $p$ enlarges, the mmWave channel has worse connectivity, thus the thresholds goes down and converges to always choosing the sub-6GHz channel.

Then we compare our optimal scheduling policy (called Age-optimal) with three other policies, including (i) always choosing the mmWave channel (called mmWave), (ii) always choosing the sub-6GHz channel (called sub-6GHz), and (iii) randomly choosing the mmWave and sub-6GHz channels with equal probability (called Random). We provide the performance of these policies for different $q$ in Fig. 5 and Fig. 6. Our optimal policy outperforms other policies.

If the two channels have a similar age performance, the benefit of the optimal policy enlarges as the mmWave channel becomes positively correlated ($q$ is larger). If the two channels have a large age performance disparity, the optimal policy is close to always choosing a single channel, and thus the benefit is obviously low. Although our theoretical results consider linear age, we also provide numerical results when the cost function is nonlinear on age by using value iteration [28]. For exponential age in Fig. 7, the gain is significantly large for all $q$; other policies have more than 2 times of average cost than the optimal policy. The numerical simulation indicates the importance of exploring optimal policy for nonlinear age cost function, which is our future research direction.

6 CONCLUSION

In this paper, we have studied age-optimal transmission scheduling for hybrid mmWave/sub-6GHz channels. For all possibly values of the channel parameters and the ON-OFF state of the mmWave channel, the optimal scheduling policy have been proven to be of a threshold-type on the age. Low complexity algorithms have been developed for finding the optimal scheduling policy. Finally, our numerical results show that the optimal policy can reduce age performance disparity, the optimal policy is close to always choosing the sub-6GHz channel.

ACKNOWLEDGEMENT

This work was funded in part through NSF grants: CNS-1901057, CNS-2007231, CNS-1618520, and CNS-1409336, and an Office of Naval Research under Grant N00014-17-1-241.

REFERENCES

7 APPENDICES

In this section we prove our main results: Theorem 1 (Section 7.2) and Theorem 2 (Section 7.3). In Section 7.1, we describe a discounted problem that helps to solve average problem (2). In Section 7.2, we introduce Proposition 1 which plays an important role in proving Theorem 1. Section 7.3 provides the proof of Theorem 2.

7.1 Preliminaries

To solve Problem (2), we introduce a discounted problem below. The objective is to solve the discounted sum of expected cost given an initial state s:

\[ J^\alpha(s) = \inf_{\pi \in \Pi} \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{i=0}^{T} E[\alpha^i \Lambda^\alpha(t)|s(0) = s], \]

(34)

where \( \alpha \in (0, 1) \) is the discount factor. We call \( J^\alpha(s) \) the value function given the initial state s. Recall that we use \( s = (\delta, l_1, l_2) \) to denote the system state, where \( \delta \) is the age value and \( l_1, l_2 \) are the state of Channel 1 and Channel 2. From Lemma 1, we only need to consider \( \pi \in \Pi^* \) instead of \( \pi \in \Pi \).

The value function \( J^\alpha(s) \) satisfies a following property:

**Lemma 3.** For any given \( \alpha \) and s, \( J^\alpha(s) < \infty. \)

**Proof.** See our technical report [25].

A policy \( \pi \) is deterministic stationary if \( \pi(t) = Z(s(t)) \) at any time t, where \( Z : S \rightarrow \Pi \) is a deterministic function. According to [34], and Lemma 3, there is a direct result for Problem (34):

**Lemma 4.** (a) The value function \( J^\alpha(s) \) satisfies the Bellman equation

\[ Q^\alpha(s, u) = \delta + \alpha \sum_{s' \in S} P_{ss'}(u) J^\alpha(s'), \]

and

\[ J^\alpha(s) = \min_{u \in U} Q^\alpha(s, u). \]

(b) There exists a deterministic stationary policy \( \mu^\alpha(s) \) that satisfies Bellman equation (35). The policy \( \mu^\alpha(s) \) solves Problem (34) for all initial state s.

(c) Assume that \( J^\alpha_0(s) = 0 \) for all s. For \( n \geq 1, J^\alpha_n \) is defined as

\[ Q^\alpha_n(s, u) = J^\alpha_n(s) = \min_{u \in U} Q^\alpha_n(s, u), \]

\[ J^\alpha_n(s) = \min_{u \in U} Q^\alpha_n(s, u), \]

then \( \lim_{n \rightarrow \infty} J^\alpha_n(s) = J^\alpha(s) \) for every s.

Also, since the cost function is linearly increasing in age, utilizing Lemma 4(c), we also have

**Lemma 5.** For all given \( l_1 \) and \( l_2, J^\alpha(\delta, l_1, l_2) \) is increasing in \( \delta. \)

**Proof.** See our technical report [25].

Since Problem (34) satisfies the properties in Lemma 4, utilizing Lemma 4 and Lemma 5, the following Lemma gives the connection between Problem (2) and Problem (34).

**Lemma 6.** (a) There exists a stationary deterministic policy that is optimal for Problem (2).

(b) There exists a value \( J^\alpha \) for all initial state s such that

\[ \lim_{\alpha \rightarrow 1} (1 - \alpha) J^\alpha(s) = J^\alpha. \]

Moreover, \( J^\alpha \) is the optimal average cost for Problem (2).

(c) For any sequence \((\alpha_n)\) of discount factors that converges to 1, there exists a subsequence \((\alpha_n^*)\) such that \( \lim_{n \rightarrow \infty} J^{\alpha_n^*} = J^\alpha \). Also, \( \mu^\alpha \) is the optimal policy for Problem 2.

**Proof.** See our technical report [25].

Lemma 6 provides the fact that: We can solve Problem (34) to achieve Problem (2). The reason is that the optimal policy of Problem (34) converges to the optimal policy of Problem (2) in a limiting scenario (as \( \alpha \rightarrow 1 \)).

7.2 Proof of Theorem 1

We begin with providing an optimal structural result of discounted policy \( \mu^\alpha. \) Then, we achieve the average optimal policy \( \mu^\alpha \) by letting \( \alpha \rightarrow 1. \)

**Definition 3.** For any discount factor \( \alpha \in (0, 1), \) the channel parameters \( p, q \in (0, 1) \) and \( d \in \{2, 3, \ldots\}, \) we define

\[ B_1(a) = \{(p, q, d) : F(p, q, d, a) \leq 0, H(p, q, d, a) \leq 0\}, \]

\[ B_2(a) = \{(p, q, d) : F(p, q, d, a) > 0, G(p, q, d, a) \leq 0\}, \]

\[ B_3(a) = \{(p, q, d) : F(p, q, d, a) > 0, G(p, q, d, a) > 0\}, \]

\[ B_4(a) = \{(p, q, d) : F(p, q, d, a) \leq 0, H(p, q, d, a) > 0\}, \]

where functions \( F(\cdot), G(\cdot), H(\cdot) : \Theta \times (0, 1) \rightarrow \mathbb{R} \) are defined as:

\[ F(p, q, d, a) = \sum_{i=0}^{d} (ap)^i - \sum_{i=0}^{d-1} a^i, \]

\[ G(p, q, d, a) = 1 + a(1-q) \sum_{i=0}^{d-1} a^i - \sum_{i=0}^{d-1} a^i, \]

(38)
Observe that all four regions $B_i(\alpha)$ converge to $B_1$ as the discount factor $\alpha \to 1$, where the regions $B_i$ are described in Definition 6. The optimal structural result of Problem (34) with a discount factor $\alpha$ is provided in the following proposition (Note that Theorem 1 can be immediately shown from Proposition 1, Lemma 6 and the convergence of the regions $B_i(\alpha)$ to $B_1$ (for $i = 1, 2, 3, 4$) as $\alpha \to 1$):

**Proposition 1.** There exists a threshold type policy $\mu^{\alpha,*}(\delta, l, 0)$ on age $\delta$ that is the solution to Problem (34) such that:

(a) If $l_1 = 0$ and $(p, q, d) \in B_1(\alpha) \cup B_2(\alpha)$, then $\mu^{\alpha,*}(\delta, l, 0)$ is non-increasing in the age $\delta$.

(b) If $l_1 = 0$ and $(p, q, d) \in B_3(\alpha) \cup B_4(\alpha)$, then $\mu^{\alpha,*}(\delta, l, 0)$ is non-decreasing in the age $\delta$.

(c) If $l_1 = 1$ and $(p, q, d) \in B_3(\alpha) \cup B_4(\alpha)$, then $\mu^{\alpha,*}(\delta, l, 0)$ is non-increasing in the age $\delta$.

(d) If $l_1 = 1$ and $(p, q, d) \in B_3(\alpha) \cup B_4(\alpha)$, then $\mu^{\alpha,*}(\delta, l, 0)$ is non-decreasing in the age $\delta$.

Since Channel 1 and Channel 2 have different delays, we are not able to show that the optimal policy is threshold type by directly observing the Bellman equation like [1]. Thus, we will use the concept of supermodularity [41, Theorem 2.8.2]. The domain of age set and decision set in the Q-function is a lattice. Given a positive $s$, the subset $[s, s + 1, \ldots] \times \{1, 2\}$ is a sublattice of $[1, 2, \ldots] \times \{1, 2\}$. Thus, if the following holds for all $\delta > s$:

$$
Q^\alpha(\delta, l, 1, 0) - Q^\alpha(\delta - 1, l, 1, 0) \\
Q^\alpha(\delta, l, 2, 0) - Q^\alpha(\delta - 1, l, 1, 0),
$$

then the Q-function $Q^\alpha(\delta, l, 0, u)$ is supermodular in $(\delta, u)$ for $\delta > s$, which means the optimal decision

$$
\mu^{\alpha,*}(\delta, l, 1, 0) = \arg\min_u Q^\alpha(\delta, l, 1, 0, u)
$$

is non-increasing in $\delta$ for $\delta \geq s$. If the inequality of (39) is inverted, then we call $Q^\alpha(\delta, l, 0, u)$ submodular in $(\delta, u)$ for $\delta > s$, and $\mu^{\alpha,*}(\delta, l, 1, 0)$ is non-decreasing in $\delta$ for $\delta \geq s$.

For ease of notations, we give Definition 4:

**Definition 4.** Given $l_1 \in \{0, 1\}$, $u \in \{1, 2\}$,

$$
L^\alpha(\delta, l, u) \triangleq Q^\alpha(\delta, l, 1, 0, u) - Q^\alpha(\delta - 1, l, 1, 0, u).
$$

Note that $L^\alpha(\delta, l, 1)$ is the left hand side of (39), and $L^\alpha(\delta, l, 2)$ is the right hand side of (39).

However, because of the mismatch of delays in our problem, most of the well-known techniques to show supermodularity (e.g., [28],[23],[19] etc) do not apply in our setting. Thus, we need a new approach to show the supermodularity. Our key idea is as follows: First, we show that $L^\alpha(\delta, l, 1, 2)$ is a constant (see Lemma 7 below), then we compare $L^\alpha(\delta, l, 1, 1)$ with the constant to check supermodularity and get Lemma 8 and Lemma 9 below. The proof of comparing $L^\alpha(\delta, l, 1, 1)$ with the constant is relegated to the proofs of Lemma 8 and Lemma 9 in our technical report [25].

Suppose that $m \triangleq \sum_{i=0}^{d-1} a^i$, and we have:

**Lemma 7.** For all $\delta \geq 2$ and $l_1 \in \{0, 1\}$, $L^\alpha(\delta, l, 1, 2) = m$.

**Proof.** See our technical report [25].

Also, we have

**Lemma 8.** (a) If $l_1 = 0$ and $(p, q, d) \in B_1(\alpha) \cup B_4(\alpha)$, then $Q^\alpha(\delta, l, 1, 0, u)$ is supermodular in $(\delta, u)$ for $\delta \geq 2$.

(b) If $l_1 = 0$ and $(p, q, d) \in B_2(\alpha) \cup B_3(\alpha)$, then $Q^{\alpha}(\delta, l, 1, 0, u)$ is submodular in $(\delta, u)$ for $\delta \geq 2$.

**Proof.** See our technical report [25].

**Lemma 9.** (a) If $l_1 = 1$ and $(p, q, d) \in B_1(\alpha) \cup B_4(\alpha)$, then there exists a positive integer $s$, such that $Q^\alpha(\delta, l, 1, 0, u)$ is supermodular in $(\delta, u)$ for $\delta > s$, and $\mu^{\alpha,*}(\delta, l, 1, 0)$ is always 1 or always 2 for all $\delta \leq s$.

(b) If $l_1 = 1$ and $(p, q, d) \in B_2(\alpha) \cup B_3(\alpha)$, then there exists a positive integer $s$, such that $Q^\alpha(\delta, l, 1, 0, u)$ is submodular in $(\delta, u)$ for $\delta > s$, and $\mu^{\alpha,*}(\delta, l, 1, 0)$ is always 1 or always 2 for all $\delta \leq s$.

**Proof.** See our technical report [25].

7.3 **Proof of Theorem 2**

For $(p, q, d) \in B_1$, we firstly prove that $\mu^\alpha(\delta, 0, 0) = 1$ and then show $\mu^\alpha(\delta, 1, 0) = 1$.

**Lemma 10.** If $(p, q, d) \in B_1 \cup B_4$, then the optimal decision at states $(\delta, 0, 0)$ for all $\delta$ are 1.

**Proof.** See our technical report [25].

On the other hand, when $l_1 = 1$, we have the following:

**Lemma 11.** If $(p, q, d) \in B_1$, then the optimal decision $\mu^\alpha(1, 1, 0) = 1$.

**Proof.** See our technical report [25].

Theorem 2(a) follows directly from Lemma 10 and Lemma 11. Due to the space limit, the proof of Theorem 2(b) when $(p, q, d) \in B_2$ is relegated to our technical report [25].