



NRL/MR/5521--20-10,212

Age of Information: Control and Estimation

SASTRY KOMPPELLA

CLEMENT KAM

*Networks & Communication Systems
Information Technology Division*

December 22, 2020

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE (DD-MM-YYYY) 22-12-2020			2. REPORT TYPE NRL Memorandum Report			3. DATES COVERED (From - To) 01-10-2018 – 30-09-2020			
4. TITLE AND SUBTITLE Age of Information: Control and Estimation						5a. CONTRACT NUMBER			
						5b. GRANT NUMBER			
						5c. PROGRAM ELEMENT NUMBER 61153N			
6. AUTHOR(S) Sastry Kompella and Clement Kam						5d. PROJECT NUMBER			
						5e. TASK NUMBER			
						5f. WORK UNIT NUMBER 1J22			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Research Laboratory 4555 Overlook Avenue, SW Washington, DC 20375-5320						8. PERFORMING ORGANIZATION REPORT NUMBER NRL/MR/5521--20-10,212			
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Naval Research Laboratory 4555 Overlook Avenue, SW Washington, DC 20375-5320						10. SPONSOR / MONITOR'S ACRONYM(S) NRL			
						11. SPONSOR / MONITOR'S REPORT NUMBER(S)			
12. DISTRIBUTION / AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A: Approved for public release; distribution is unlimited.									
13. SUPPLEMENTARY NOTES									
14. ABSTRACT The purpose of this work is to study the impact of the timeliness of information exchange on application performance, by developing a deeper analytical understanding of the AoI metric over a varied set of network instantiations. We developed new tools for the estimation and reconstruction of the real-time process being observed from across a network, and studied the impact of timeliness of information exchange on the optimal reconstruction process in a novel way using the AoI metric. Additionally, our study also included information freshness for multiple competing sources and multiple transmission hops.									
15. SUBJECT TERMS Age of Information (AoI) Infomation freshness Latency Remote signal estimation Signal sampling Stochastic hybrid systems (SHS)									
16. SECURITY CLASSIFICATION OF:						17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Sastry Kompella	
a. REPORT Unclassified Unlimited	b. ABSTRACT Unclassified Unlimited	c. THIS PAGE Unclassified Unlimited	Unclassified Unlimited		21			19b. TELEPHONE NUMBER (include area code) (202) 767-5098	

This page intentionally left blank.

CONTENTS

EXECUTIVE SUMMARY	E-1
1. INTRODUCTION.....	1
2. TECHNICAL OBJECTIVE.....	2
3. AGE OF INFORMATION FUNDAMENTALS.....	2
4. REMOTE SIGNAL ESTIMATION	3
4.1 Effective Age Metric.....	3
4.2 Estimation of an N -State Markov Source.....	5
4.3 Learning to Sample	8
4.4 Age of Incorrect Information	10
4.5 Effect of Channel State Information (CSI)	10
5. REAL-TIME STATUS UPDATING WITH MULTIPLE SOURCES.....	11
5.1 Information Freshness over an Interference Channel	11
5.2 Multi-source Systems.....	13
6. MULTI-HOP WIRELESS NETWORKS	14
7. FUTURE RESEARCH DIRECTION.....	15
REFERENCES	16

FIGURES

Fig. 1 – Evolution of the <i>age</i> metric.....	3
Fig. 2 – (a) Age, and (b) Error % versus p	4
Fig. 3 – Effective Age Metrics: (a) Sampling Age, and (b) Cumulative Marginal Error.....	5
Fig. 4 – N -state Markov Chain.	6
Fig. 5 – Age versus p , $\mu = 0.1$, 7 states.....	6
Fig. 6 – MSE vs p , 7 states: (a) $\mu = 0.1$ (b) $\mu = 0.5$ (c) $\mu = 1.0$	6
Fig. 7 – Full State versus BLR, $p=0.1$: (a) Age vs. #bits per state (b) MSE vs. #bits per state.....	7
Fig. 8 – Exponential server with $\mu = 1$ (averaged over 1000 simulations): (a) Average AoI vs. Time (b) % Error vs. Time.....	8
Fig. 9 – Exponential server with $\mu = 1$ followed by $\mu = 0.5$: (a) Average AoI vs. Time (b) %Error vs. Time.....	9
Fig. 10 – AoI versus Time for a log-normal server with $\sigma = 2$ (averaged over 1000 simulations).....	9
Fig. 11 – Geometrically distributed delay, $p_s = 0.5$: (a) Real-time error, (b) Age of Information, (c) Age of Incorrect Information.....	10
Fig. 12 – Age vs. p_{22} under various types of CSI for (a) $p_{11} = 0.1$, and (b) $p_{11} = 0.9$	11
Fig. 13 – Information updates over an interference channel.	11
Fig. 14 – NE for various (bit rate, BER) pairs: (a) Transmission power, (b) AoI, and (c) Total cost.....	13
Fig. 15 – Multi-source System.	13
Fig. 16 – Queues in tandem.	14
Fig. 17 – Average age vs. λ , $\mu = 1$, queue capacity of 1. Simulation minima indicated by Δ	15
Fig. 18 – Average age vs. λ , $\mu = 1$: (a) infinite queue capacity, and (b) k queue capacity.....	15

TABLES

No table of figures entries found.

EXECUTIVE SUMMARY

The concept of Age of Information (AoI) is useful in any communication system where the receiver has an interest in fresh (most recently generated) information. This is indeed the case in numerous applications of wireless systems which require the transmission of status updates between users, such as in sensor networks, environment monitoring, and more importantly Navy/Marine Corps situational awareness type applications. Traditional network metrics of throughput and delay are inadequate for describing the performance of these status monitoring applications. This report describes our work as part of a three-year NRL base program funded project effort titled “Age of Information: Control and Estimation,” where we addressed the impact of the timeliness of information exchange on application performance, by developing an analytical understanding of the Age of Information metric over a varied set of network instantiations.

As a metric for information freshness, we defined the instantaneous *age* of an update message at the time of observation as the current (observation) time minus the time at which the observed state was generated. Therefore, minimizing the age directly describes the objective of achieving timely updates at the monitoring station in a way that traditional metrics such as throughput and delay do not. At the beginning of this project, our understanding of the AoI metric was based on modeling simple networks as different types of queues, and with various arrival/departure processes. In doing so, we had not really considered the actual real-time process (signal) that the sources are observing, just that these observations are packetized at discrete intervals of time into update messages and delivered over the network to a remote monitor. Our hypothesis for this project was that the performance of any update monitoring system depends on actually recreating the observed process by the source at the destination with high fidelity, which requires prediction / estimation of the observed signal based on the update messages that have been successfully received by the monitoring station. It is easy to see that such estimation when performed intelligently will have an impact on the age of information, but more importantly, the age will have a significant impact on the estimation process.

The purpose of this work is to two-fold. First, as described earlier, we wanted to study the impact of signal sampling and the *age* of the update messages (sample measurements) generated by one or more sources on the estimation of the observed real-time process at the monitoring stations. We learned that while AoI indeed has a role to play, optimizing the AoI metric does not always guarantee optimal signal reconstruction at the monitoring station. A second dimension to our work is to understand how information freshness is impacted by competing sources, and when updates have to travel beyond a single hop. Better understanding of the relationship between these factors has given us more insights towards developing better update-monitoring systems by optimizing for signal estimation / prediction, thereby, enhancing situational awareness.

This page intentionally left blank.

AGE OF INFORMATION: CONTROL AND ESTIMATION

1. INTRODUCTION

The concept of Age of Information is useful in any communications system where the receiver has an interest in fresh information. This is indeed the case in numerous applications of wireless systems that require the transmission of status updates between nodes, such as sensor networks, situational awareness applications, and environment monitoring. Traditional network metrics of throughput and delay are inadequate for describing the performance of these status monitoring applications. This is because the delay for a particular packet may be small, but if its transmission occurred a long time ago, the information, as observed at the current time, is no longer fresh. As another example, throughput may be high, such that packets arrive very frequently at the receiver. However, since the queues are usually saturated at high throughput, the packets that arrive at the destination are typically generated a long time ago but delayed in a queue (at the source or relay nodes). This results in even recently received packets no longer being fresh. Therefore, we realized in the last few years that a different metric such as Age of Information (AoI) is needed to convey the freshness of information at the receiver.

This report describes our work as part of a three-year NRL base program funded project effort titled “Age of Information: Control and Estimation,” where we addressed the impact of the timeliness of information exchange on application performance, by developing a deeper analytical understanding of the AoI metric over a varied set of network instantiations. In our prior work leading up to this project, we had developed a rudimentary understanding of the AoI metric over simple network settings, and showed that minimizing it results in achieving timely updating in a way that traditional metrics such as throughput and delay do not. Our research on the age metric had focused on modeling communication between a source and a monitoring station, using different types of queues, with various arrival/departure processes, number of servers, and queue capacities. In particular, this communication was modeled as an M/M/1, M/M/2 and in the extreme case, the M/M/ ∞ system (nomenclature described in detail later in Section 3). Analyzing for age showed that when the number of servers is finite, there is an optimum packet generation rate for which the average ages is minimized. Also, the availability of more network resources results in the lowering of the average age. However, the tradeoff is that more packets will arrive out of order, and those packets that do not update the latest status are essentially useless and end up wasting network resources. Additionally, it was also shown that deterministic arrival and departure processes achieve a lower average age than memoryless processes. Further work we performed on packet queue management along with the use of packet deadlines added a new dimension to optimizing the age of information.

While traditional networking focuses on traffic engineering of incoming data flows, a key distinction of the AoI research is the focus on generating data for status updating, or stated differently, sampling a signal. Specifically, an important part of the system that had yet to be considered is the structure of the process that we are observing in a discrete-time fashion, which will have serious implications on the nature of transmission / estimation strategies that should be employed. We realized early on that understanding the complex interaction between sampling and the delay induced by the network is fundamental to optimizing information freshness. Therefore, as part of this project we wanted to study the impact of the age metric, sampling rate and sampling policy on application performance, specifically signal estimation at the monitoring station. We have also studied mechanisms for controlling how these samples flow into the network, through deadline mechanisms or controlling the queue capacity. We were able to extend our knowledge of the sampling-delay effect on information freshness in a complex network of systems with heterogeneous goals.

Another recent branch of the AoI research has focused on alternative freshness metrics related to age, in cases where the actual freshness is not the ultimate objective, but rather the value of freshness to the user.

One such example is the concept of age-penalty functions, which allows for different levels of tolerance of information staleness, allowing for optimization of scenario- or user-specific freshness needs. Another is the pursuit of an effective age metric for real-time remote estimation. We demonstrated as part of this project that the age metric itself does not completely characterize the remote estimation performance, and proposed effective age metrics that sufficiently described the performance for the signal in question (e.g., a Markovian source). The search for an effective age metric that is generalizable to any signal model is ongoing. While these efforts have started to consider objectives beyond the age metric, there is still a lack of understanding of the relationship between age and other metrics important to the particular mission, such as accuracy and completeness of data, which will be our focus in the future.

Finally, we wanted to understand how information freshness is impacted by competing sources, and when updates have to travel beyond a single hop, which was the focus of the last year of the project. Better understanding of the relationship between all these factors has given us more insights towards developing better update-monitoring systems by optimizing for signal estimation / prediction, thereby, enhancing situational awareness.

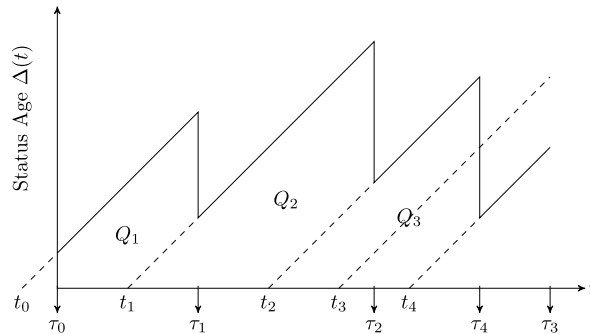
We believe our work in this project has developed novel communication / networking techniques that improve our understanding of information latency in communication systems and will ultimately enhance the performance of Navy/Marine Corps applications, general DOD operations, and disaster relief / emergency applications. These techniques comprise of algorithmic methods of network control and evaluation, and are designed to adapt to the time variability of tactical networks.

2. TECHNICAL OBJECTIVE

The purpose of this work is to study the impact of the timeliness of information exchange on application performance by developing a deeper analytical understanding of the AoI metric over a varied set of network instantiations. We developed new tools for the estimation and reconstruction of the real-time process being observed from across a network, and studied the impact of timeliness of information exchange on the optimal reconstruction process in a novel way using the AoI metric. We also developed newer Effective Age metrics such as Age of Incorrect Information, where the original AoI metric proved insufficient, and studied the impact of timely delivery of channel state information on the reconstruction process. Additionally, our objective for this project also included the study of these new dimensions of information freshness in the presence of multiple sources and multiple transmission hops.

3. AGE OF INFORMATION FUNDAMENTALS

The age metric is defined as $\Delta(t) = t - u(t)$, where $u(t)$ is the time stamp of the most recently received packet by the monitor. Considering a simple system of a first-come, first-served queue, Fig. 1 shows how the age metric typically evolves over time. We see from Fig. 1 that a packet is received at the monitor at time τ_0 , and at time 0, the age is equal to $-\tau_0$. As time progresses, the information ages linearly until the next packet is received at time τ_1 , which has an age equal to $\tau_1 - \tau_1$. Continuing this process, we observe that the function exhibits a “saw-tooth” pattern that is characteristic of the age function.

Fig. 1 – Evolution of the *age* metric.

Typical AoI studies model the system as memoryless and have a source that transmits packets to a monitor through an M/M/1/K queue, where the last entry in the Kendall notation describes a total capacity of $(K - 1)$ packets in the queue and one packet in service. In most cases, an arriving packet that encounters a full capacity system never enters the system and is dropped. We refer to the time between packet generations as the interarrival time $X_i = (t_i - t_{i-1})$, $i = 1, 2, 3, \dots$. The interarrival times are modeled as random if the source does not have control over the exact times at which it can transmit updates. In that case, typically, the X_i 's are modeled as i.i.d. exponential random variables with rate λ . We call the time spent in the server by packet k the service time S_k , $k = 1, 2, \dots$. The service time S_k is typically modeled as exponential with rate μ , and all the S_k 's are i.i.d. and independent of the X_i 's. The total time spent in the system from arrival to service is given by T_k , $k = 1, 2, \dots$, where $T_k = W_k + S_k$, with W_k being the time spent waiting in the queue. We also define the inter-departure time Y_k as the time between the instants of complete service for the $(k - 1)$ st packet served and the k th packet served. This will be useful in the computation of the average age.

The time-average age can be obtained using a graphical approach to compute the area under the saw-tooth curve, specifically by computing the average area of a trapezoidal shape associated with each packet updating the age. This is a non-trivial computation even for simple queueing systems (e.g., M/M/1 queue), since it involves computing the expected value of the product of the packet inter-arrival time at the source and the system time, which are not independent. Even for slightly more complex networks, the average age analysis is a difficult task.

Specific aspects of the work accomplished in this project include the following interrelated research thrusts described in the rest of the document, which has helped form the basis of our deeper understanding of information latency in complex systems.

4. REMOTE SIGNAL ESTIMATION

One of the core aspects of this project revolves around optimal reconstruction of a signal / process that is being observed by the source node(s) at the receiver of the update monitoring system. Specifically, as part of this project, our goal was to estimate the observed signal at the receiver with the least possible error, and we wanted to understand the role that information freshness, and thereby age of information, plays in this regard.

4.1 Effective Age Metric

It had been shown earlier that for the remote estimation problem, the sampling strategy that minimizes the age of information (AoI) does not minimize the estimation error. Therefore, we worked on obtaining an

alternative metric, called effective age, for which a lower effective age necessarily yields a lower estimation error. The problem we considered as our basis for developing an effective age metric is the remote estimation of a Markov source, in which samples of the source signal are transmitted over a delay channel to be estimated at a destination.

We considered a slotted-time system in which a source generates samples just before a chosen slot to be sent over an error-free delay channel. The transmission delay for the i th packet, S_i , is an independent geometric random variable with probability of success in each slot equal to $\mu \in (0,1]$. The packet is fully received at the destination at the end of S_i . The source knows when the system is busy transmitting, and thus there is no need to sample during a busy period to ensure fresh samples. Following a completed transmission S_i , the source decides how many slots to wait before generating a new sample to begin transmission immediately. This is referred to as the “generate-at-will” model. At the source node (transmitter), there is a discrete Markov process X_0, X_1, \dots , where $X_n \in \{0, 1\}$, that we would like to estimate in real time at the destination. The duration of each state X_n is equal to a transmission slot length, and if a state transitions occurs, it happens just prior to the sampling decision before each slot. The probability of changing states is equal to p .

We compared two sampling strategies. The first is the *zero-wait* strategy, in which a sample is generated immediately after the previous sample was received at the destination. The second strategy is the *sample-at-change* strategy, in which a sample is generated only when the state changes from the previously received state. We analyzed the AoI and estimation error for this system, and proved, as shown in Fig. 2, that while the first strategy yields the lower age, the second strategy yields the lower estimation error.

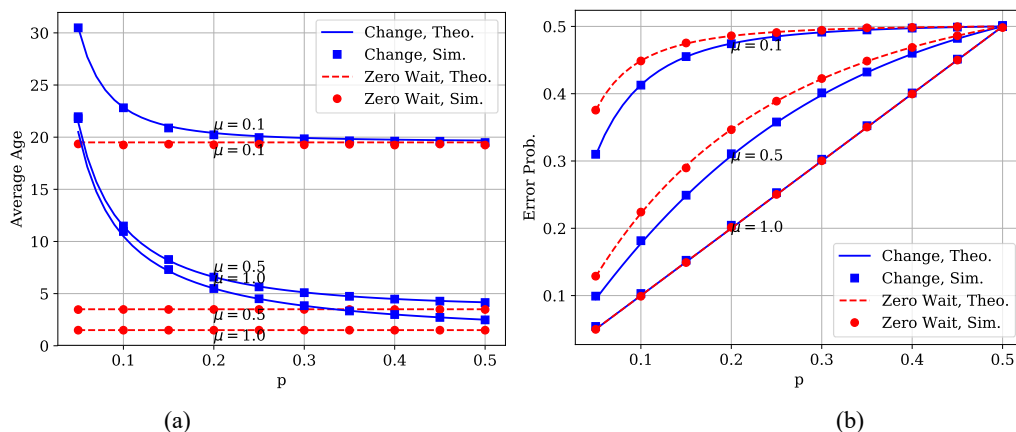


Fig. 2 – (a) Age, and (b) Error % versus p .

Insights into the sampling strategy and its impact on AoI and error performance enabled us to develop and study effective age metrics. Specifically, the sample-at-change strategy uses the knowledge of the source signal to choose the sample at the exact moment the state changes, providing the freshest information with respect to each change of the signal. With this insight, we presented two candidate effective age metrics: sampling age and cumulative marginal error as shown in Fig. 3. The sampling age tracks the age of the samples relative to the ideal sampling time, which in the two-state Markov source model is the time at which the state changes. The cumulative marginal error tracks the total error from the reception of the latest sample to the current time.

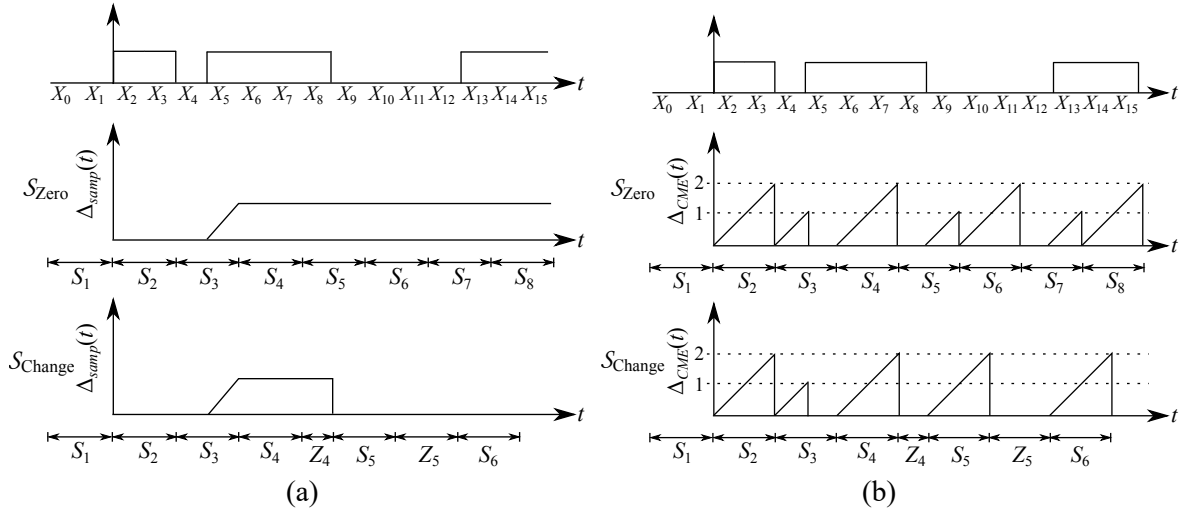


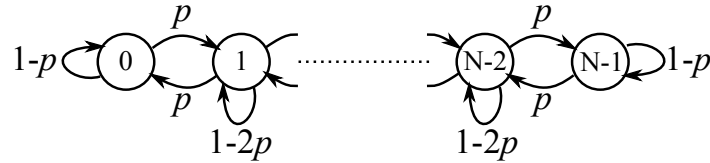
Fig. 3 – Effective Age Metrics: (a) Sampling Age, and (b) Cumulative Marginal Error.

We observed that both metrics demonstrate a monotonic relationship with respect to estimation error for the Markov source system, and in [1] we detail some limitations of the metrics as well as benefits, such as the CME being a way to blend the two goals of minimizing age and error.

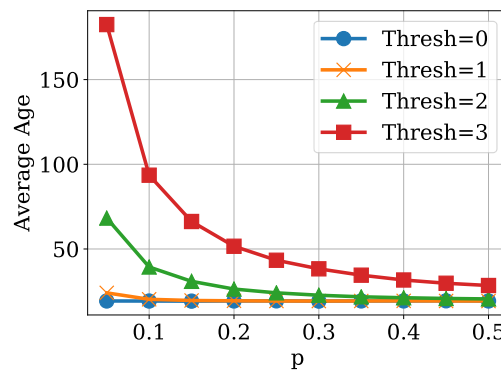
4.2 Estimation of an N -State Markov Source

Continuing with the concept of effective age, we specifically wanted to understand, in the case of a Markov source, why the sampling pattern that minimizes the average age or average peak age does not necessarily minimize the estimation error. We reasoned that this apparent disparity is due to the fact that the value of the age is determined by two separate factors. One is the processing/transmission delay (including possible queueing delay). The other is the sampling pattern. Thus, a first concern was to consider modifying the “physical” or “actual” age so that its average value minimization (when we choose the sampling pattern) also achieves minimization of the prediction/estimation error. A second concern was to modify the definition of the physical age so as to capture more meaningfully the information content of the process. For example, for a process with strong correlation properties, the physical age is not a good indicator of its prediction error. These concerns motivated our efforts to propose an “effective” age, the average value of which would correspond to the size of the error, and would capture the information structure (or content) of the process. This is of course a rather tall order, and in [2], we only began to chart out a path towards these goals.

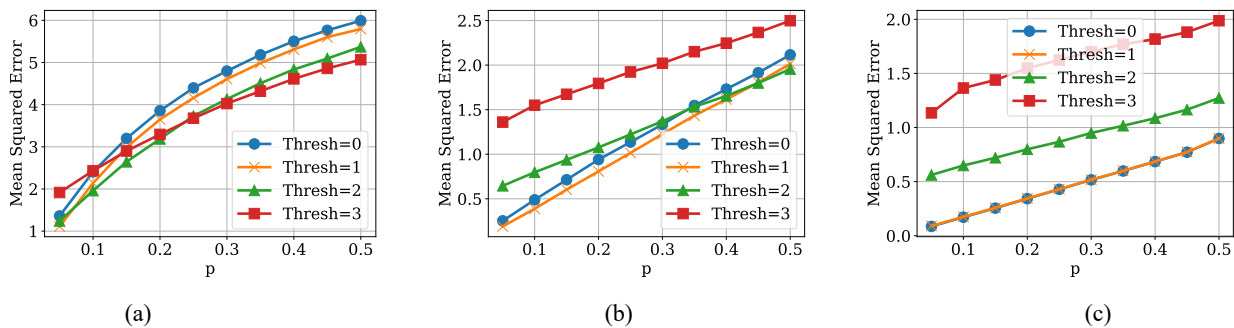
We started with a two-part effort. First, we extended some of our earlier work on sampling patterns for the simplest of signals, namely a 2-state Markov Chain to a generalized N -state Markov Chain, as shown in Fig. 4. The signal that is being sampled, therefore, is a discrete-time Markov source in which the time step is equal to a time slot length of the communication system. The Markov process is again given by X_n , where the system transitions from X_n to X_{n+1} with a probability $\Pr(X_{n+1} | X_n)$. Although this model can be kept general, for the purpose of this work, we consider a special case in which we have an N -state birth-death type of Markov chain, where transitions only occur between state $(i-1)$ and i for $1 \leq i \leq N-1$, and self-transitions. Specifically, we consider the transition probabilities to be $\Pr(X_{n+1} = i | X_n = i-1) = \Pr(X_{n+1} = i-1 | X_n = i) = p$, for $1 \leq i \leq N-1$.

Fig. 4 – N -state Markov Chain.

We again considered the zero-wait sampling pattern at the source and compared the average age to that of a *threshold-like* sampling policy that generates a sample only when the state transitions to another that is some distance away from the last sampled state. Our error parameter is the squared error in estimating the current value of the signal based on the most recent sample. We simulated a 7-state Markov source to evaluate how the optimal threshold value changes with p . We ran each simulation so that there would be on average 200 state changes, and we averaged the age over time and over 1000 simulation runs. In Fig. 5, we have plotted the average age as a function of p for threshold values 0,1,2,3. As expected, the age is larger for larger thresholds, since the source waits for a larger change before sampling.

Fig. 5 – Age versus p , $\mu = 0.1$, 7 states.

In Fig. 6, we show the MSE vs. p for the various threshold values and various values of μ , which is the independent geometric random variable that signifies the transmission delay for the packet in service. Clearly, the MSE does not follow the same trend or relative ordering between the threshold values as the age. We observe in Fig. 6(a) that for a threshold of 0, the MSE is less than when the threshold is 3 for $p \leq 0.1$, but it has the highest MSE for $p \geq 0.15$. For a threshold of 1, the MSE is lowest for $p = 0.05$, but it increases rapidly with p relative to the cases where the threshold is 2 or 3.

Fig. 6 – MSE vs p , 7 states: (a) $\mu = 0.1$ (b) $\mu = 0.5$ (c) $\mu = 1.0$.

Setting the threshold to 2 yields a lower MSE for $p \leq 0.2$, but for $p \geq 0.25$, the MSE is lowest when the threshold is equal to 3. Thus, for a relatively low $\mu = 0.1$, a higher threshold than 1 should be used in most

cases. This is probably due to the likelihood of a larger change between successful receptions, and that there is more value in transmitting a sample with a larger change when receptions occur less often. Fig. 6 show the MSE for $\mu = 0.5$ and 1, respectively. We observe that as the transmission delay gets smaller, it is better to use a lower threshold since there is not much of a penalty to sampling and transmitting often.

In the second part, we took a somewhat different approach where we focused on the micro-scale of the process values by considering “bit-by-bit” the encoded values of the process and consider a scheme of sampling (i.e., of what to transmit) that is akin to a threshold-like policy but focuses on a detailed view of the encoding properties. Our take on this approach is somewhat analogous to block-fading in that it assumes that during each “clock”- time or slot, only one bit of the encoded symbol is transmitted. We believe that this restrictive model can be relaxed at some complexity cost, but it does offer some new insights. We believe that the main value of our contribution consists of the key ideas behind our models and not so much the numerical results.

Similar to the Markov source model described earlier, we have a discrete Markov process X_0, X_1, \dots , but each time step is equal to the transmission time of a single bit, and each state is represented by m bits, in which X_n is one of $2m$ discrete states. Therefore, here we only considered Markov chains where the number of states is a power of 2. The transition probability from state i to state j is given by $\Pr(X_{n+1} = i | X_n = j)$.

We simulated the bit-by-bit communication system for the various sampling strategies. We started with a comparison between continuous transmission schemes with and without bit-level replacement. In Fig. 7, we plotted the age and the MSE vs. the number of bits per state m , so that the number of states is increasing by a factor of 2. The value of p for the Markov chain is set to 0.1. Our focus was the bit-level replacement (BLR) scheme, in which we observe the state at each time slot, and if the bits of the partial state that have been sent match with the corresponding bits of the current state (mid-transmission of a full state), we send the remaining bits of the current state and refresh the time stamp with the current time. If there is no match between the current state and what has been sent, the remainder of the previous state completes transmission with the time stamp unchanged. More details regarding the BLR scheme can be found in [2]. We observed that the BLR schemes outperform the full state scheme in terms of age and MSE, and that the relative ordering is preserved from the age to MSE. For this error-free, deterministic delay system, the relative ordering was preserved, such that lower age implies lower prediction error. The optimal bit encoding does the best, but we are unable to generate results beyond 4 bits per state due to the computational complexity of the recursive algorithm.

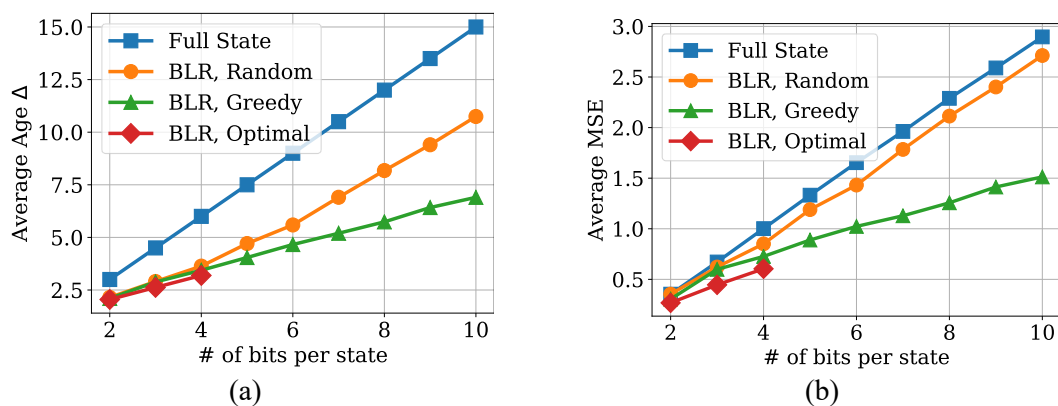


Fig. 7 – Full State versus BLR, $p=0.1$: (a) Age vs. #bits per state (b) MSE vs. #bits per state.

4.3 Learning to Sample

So far, we focused our work on estimating the observed signal at the receiver with the least possible error, and we wanted to understand the role that information freshness, and thereby age of information, plays in this regard. In this section, our focus is on a different but related aspect of information freshness: more specifically, with the knowledge of the system model [3].

Most AoI studies typically focus on models from queueing theory with some known distribution, but in practice, these models can be inadequate in describing the actual system, and/or the solutions are unable to adapt to dynamic systems. Our goal was to minimize AoI for the case in which we assume no knowledge of the system we are transmitting through, treating it as an unknown black box. We wanted to observe the age at the output of the black box and learn how to sample to minimize the age. We initially considered the special case of a lossless single-server system with an unknown service distribution. Our overall approach was to observe the age at the last packet reception, and decide how long to wait before taking a new sample at the source and transmitting. We compare two approaches: 1) estimating a parameter of the system and applying it to a parameterized model to determine the waiting time, and 2) a model-free approach that learns the waiting time based on observing the age.

For our first approach, we derived the optimal policy for an exponential server as a function of the service rate, and then at each step, we updated our service rate estimate and used it to give us the optimal waiting time. We called this the Exponential Model-based Policy (EMP). We showed that this adapts well when the server is exponential as shown in Fig. 8, and can even perform well under abrupt changes in the service rate, as shown in Fig. 9. However, when the service distribution is not exponential, this approach fails to learn the optimal waiting time. For comparison, in Fig. 8 and Fig. 9, we simulated the optimal threshold policies that have complete knowledge of the distribution of each system, as a lower bound on the age. More information is available in [3].

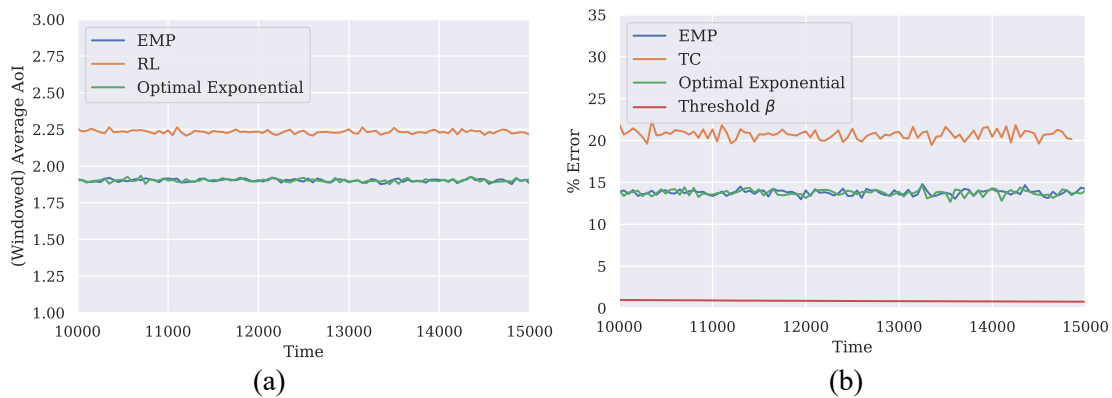


Fig. 8 – Exponential server with $\mu = 1$ (averaged over 1000 simulations): (a) Average AoI vs. Time (b) % Error vs. Time.

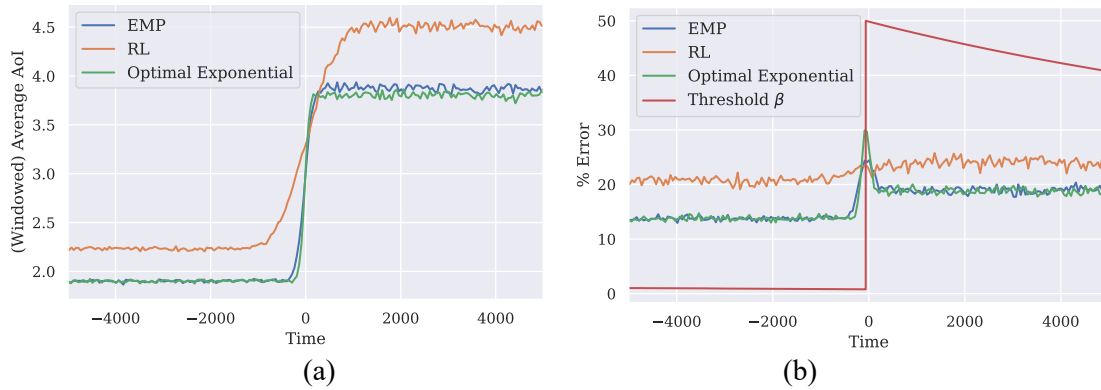


Fig. 9 – Exponential server with $\mu = 1$ followed by $\mu = 0.5$: (a) Average AoI vs. Time (b) %Error vs. Time.

In order to deal with this shortcoming, we considered a distribution-free approach based on reinforcement learning (RL), which decides actions based on a system state to optimize a reward. Further improvements in these actions can be learned by interacting with the environment. There have been a number of recent breakthroughs in machine learning and artificial intelligence using reinforcement learning, from playing board games and video games to training robots, and there are even a few recent papers on applying reinforcement learning to optimizing age of information. These current approaches consider a limited discrete action space (e.g., choosing whether to sample or not, which source to sample from), for which it is possible to tabulate the value of each action, which is a relatively more stable learning process.

In our approach, we considered a continuous action space, in which we decide the waiting time (continuous time) based on the age. Typically in reinforcement learning, the goal is to learn the action value function $Q(s,a)$, which is the expected return for taking action a in state s . The specific RL approach we use is *one-step Sarsa*, which is an on-policy temporal difference (TD) control method for updating $Q(s,a)$ based on the rewards and state transitions. We combine this Sarsa approach with *tile coding*, which allows us to approximate the action value function $Q(s,a)$ over the continuous action space.

Simulation results in Fig. 10 are shown for a log-normal distributed (given by $Y = e^{\sigma X} / E[e^{\sigma X}]$, where X is a zero-mean Gaussian random variable with unit variance) service time with parameter $\sigma = 2$. The EMP yields higher age, and has been shown in [3] that it is 116% greater than the optimal threshold policy, compared to 19% for the RL policy over the whole simulation.

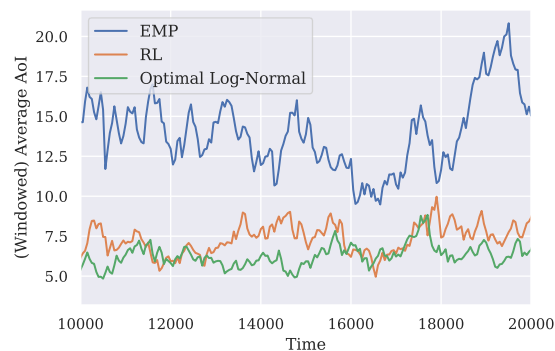


Fig. 10 – AoI versus Time for a log-normal server with $\sigma = 2$ (averaged over 1000 simulations).

4.4 Age of Incorrect Information

For monitoring applications, the Age of Information (AoI) metric has been the primary focus of recent research, but as suggested before, the problem of real-time or remote estimation is closely related. In previous works, we and others have shown Age of Information to be insufficient for minimizing remote estimation error, but recently a metric known as Age of Incorrect Information (AoII) was proposed that characterizes the cost of a monitor being in an erroneous state over time. In this section, we describe our study of the AoII metric in the simple context of monitoring a symmetric binary information source over a delay system with feedback.

Our approach was to determine the best sampling policy for the following remote monitoring/estimation performance metrics: 1) real-time error, 2) Age of Information, and 3) Age of Incorrect Information. For each performance metric, we formulated the infinite horizon, discrete time Markov Decision Process (MDP), for which we specified the *state* space S , the *action* space A of whether to transmit or remain idle when the channel is idle, the *cost* function $g(s)$, $s \in S$, and the state *transition probabilities* $P(s_{n+1} | s_n, u(s_n))$, $u(s) \in A$. For each metric, once we formulated the optimal sampling problem as an MDP, we applied a dynamic programming algorithm to compute the optimal performance and policy.

Again, we simulated the system for two sampling policies: sample-at-change and zero-wait, and we observed which policy coincides with the optimal policy for each metric. For a variety of delay distributions and AoII penalty functions, one of which is shown in Fig. 11, we observed that the optimal policy for the probability of error and for AoII is the sample-at-change policy, whereas the optimal policy for AoI is a threshold policy. More information is available in [4].

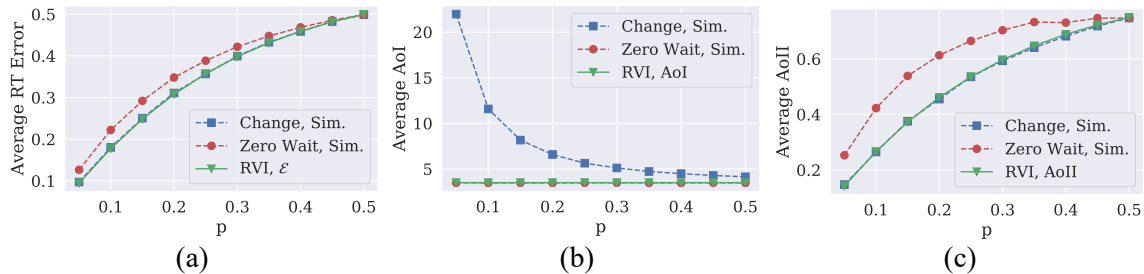


Fig. 11 – Geometrically distributed delay, $p_s = 0.5$: (a) Real-time error, (b) Age of Information, (c) Age of Incorrect Information.

4.5 Effect of Channel State Information (CSI)

While most available results on age are limited to channels that can be characterized by independent service times, our results here are for a Markov channel that is characterized by dependent service times, using the knowledge of CSI, which can be exploited to improve system performance. For example, when CSI (such as a measured/estimated channel fading level) is known for some period of time, an appropriate amount of each resource (e.g., power, time, and bandwidth) can be allocated to the users of the channel during that time period. The following cases are considered: unknown CSI, perfect CSI, erroneous CSI, and delayed CSI.

Our evaluation of the age metric for various types of CSI is shown in Fig. 12. In our evaluation, packets arrive at the source according to a Poisson process with rate λ , and then are selected for possible transmission over the Markov channel, with two states: state 1 is the good state (requiring one time slot for successful transmission of a packet) and state 2 is the bad state (requiring between 2 and k time slots for successful transmission of a packet, $k \geq 2$). Each time slot lasts t seconds. Probability p_{ij} is the

probability of transitioning from state i to state j , and π_i is the stationary probability that a time slot is in state i , where $i, j \in \{1, 2\}$.

In the numerical evaluation, we let $\lambda = 1$, $t = 1$, and $k = 2$, and plot the age metric $E[A]$ vs p_{22} for $p_{11} = 0.1$ and 0.9 . The first arrival in a time slot is selected for transmission in the next time slot, if the next slot is not occupied by a previous transmission. The curves for the unknown CSI and perfect CSI cases are obtained from the closed form results in Theorems 1 and 2 (which also agree well with simulation results) described in [6]. The other cases of CSI are more complex to analyze mathematically, and thus we rely on simulation to obtain their results.

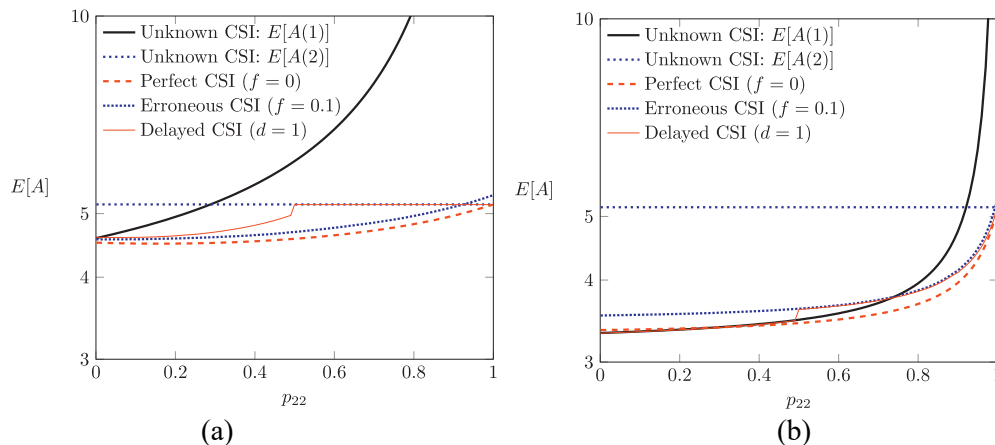


Fig. 12 – Age vs. p_{22} under various types of CSI for (a) $p_{11} = 0.1$, and (b) $p_{11} = 0.9$.

5. REAL-TIME STATUS UPDATING WITH MULTIPLE SOURCES

In this section, we describe our work on communication problems in which more than one source desires to maintain its information freshness with their intended destinations.

5.1 Information Freshness over an Interference Channel

An interference channel (IC) models the simultaneous transmissions of N independent transmitter-receiver pairs via a common (usually wireless) channel. Simultaneous transmissions can cause mutual interference. The N -user IC has N principal links and $N(N - 1)$ interference links. The capacity rate region for the general IC remains unknown. Many wireless and wireline transmissions (e.g., wireless device-to-device communications and wireline communications having cross talk between two twisted pair cables) can be modeled as ICs. In this work, we considered the transmissions over the IC, in which each transmitter desires to maintain the freshness of the information updates it sends to its intended receiver, as shown in Fig. 13.

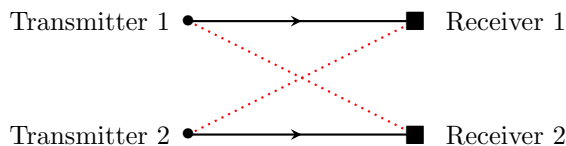


Fig. 13 – Information updates over an interference channel.

Most research on the IC has focused on the case of two transmitter-receiver pairs. Despite many years of research effort, the capacity rate region of even the 2-user IC remains unknown. The IC can be seen as a conflict/competitive situation, and it is usually not clear how to choose an acceptable operating point for users. One of the best ways to deal with this conflict is to use tools from game theory. In this work, therefore, we considered the case of two transmitter-receiver pair, and model the 2-user IC problem as a two-player “interference” game, where each player i is a transmitter desiring to maintain the freshness of the updates to its receiver, where $i \in \{1, 2\}$.

There are several well-known game strategies. For example, if the players act simultaneously, the solution is called the Nash strategy or Nash equilibrium (NE), which refers to the point at which neither player could improve its performance by unilaterally changing its strategy while the other player’s strategy remains unchanged. Alternatively, if one player (the leader) chooses its strategy and the other player (the follower) reacts by choosing its own strategy, the solution then leads to the so-called Stackelberg strategy or Stackelberg equilibrium (SE), which is usually more appropriate for military settings. In this work, we considered both the Nash and Stackelberg strategies.

Our goal was to allow each player to choose a transmission power level to minimize its total cost function, which incorporates both the information staleness (i.e., opposite of freshness) and operational cost for that player. A player’s information staleness is represented by the AoI metric as observed by its receiver, whereas the cost incurred by a player is proportional to the power transmitted by that player. Recall that the instantaneous age of an update process at time t refers to the duration of the interval from the timestamp of the most recently received update to t . Our communication -- interference model formed a continuous game, i.e., each player chose a power level as a strategy from a continuous set. We obtained the closed-form expressions for the Nash and Stackelberg strategies for the game, as well as characterized their properties. We showed that our Nash and Stackelberg strategies, along with the use of error control coding, can be used to generate desirable operating points for users in competitive/conflict situations. In particular, we showed that error control coding is beneficial when the goal is to minimize the total cost that reflects both the information freshness and the power cost by allowing higher transmission rates (despite with higher BER) over the IC. These results are described in Fig. 14.

Our two-player game model consisted of two transmitter-receiver pairs (i, i) and (j, j) , where $i, j \in \{1, 2\}$ and $i \neq j$ (see Fig. 13). The power levels transmitted by i and j are denoted by p_i and p_j , respectively. Transmitter i aims to minimize its total cost $v_i(p_i, p_j) = A_i(p_i, p_j) + w_i p_i$, where $A_i(p_i, p_j)$ is the age as observed by receiver i , and $w_i p_i$ is the power cost of transmitter i , where w_i is its *unit* power cost.

In Fig. 14, we considered the case of transmission power levels at NE. For given values of BER pair $P_e(i)$ and $P_e(j)$, the bit rates of the transmitters are given by $r_i = c \text{SIR}_i(p_i^*, p_j^*)$ and $r_j = c \text{SIR}_j(p_i^*, p_j^*)$, where c is the constant associated with binary PSK transmission, and is described in [5]. We show the results for $\text{BER} \in \{10^{-2}, 10^{-3}, 10^{-4}\}$ and the corresponding bit rates, i.e., we have three (r, P_e) pairs. Finally, r_{ECC} is the error correction coded transmission rate. Results show that we can find the required coding rate that is appropriate for a given channel bit error rate in order to minimize the total cost $v_i(p_i, p_j)$.

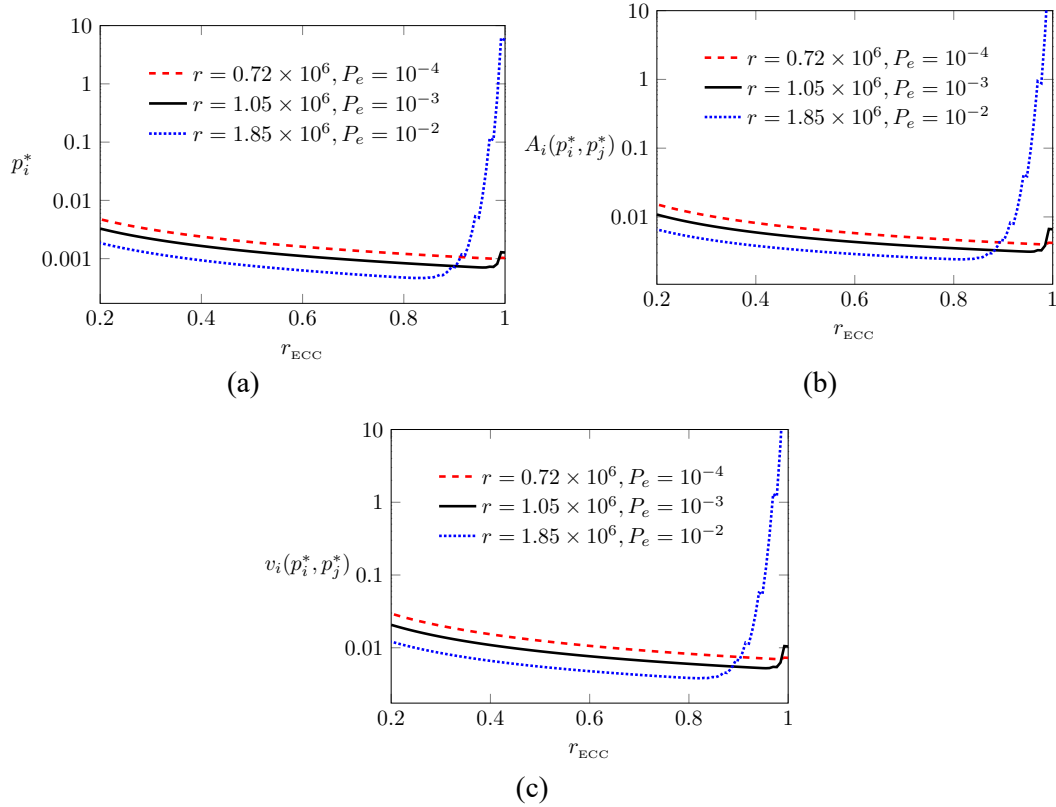


Fig. 14 – NE for various (bit rate, BER) pairs: (a) Transmission power, (b) AoI, and (c) Total cost.

5.2 Multi-source Systems

When we extend the information freshness problem space to multi-source systems, things get complicated. In this section, we describe the problem of minimizing the age of information in a multi-source system, where sources communicate their update packets to a destination via a channel with random delay [7], as shown in Fig. 15. Due to interference, only one source can be scheduled at a time.

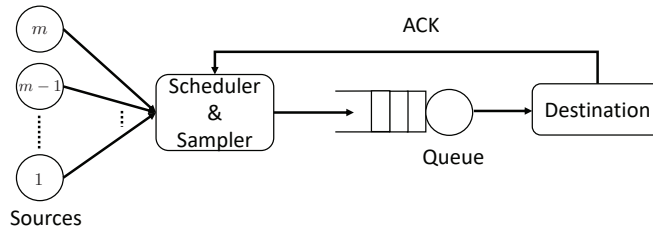


Fig. 15 – Multi-source System.

We considered the problem of finding a decision policy that controls the packet sampling times and schedules source transmissions to minimize the total average peak age (TaPA) and the total average age (TaA) of the sources. Our investigation of this problem resulted in an important separation principle: the optimal scheduling strategy and the optimal sampling strategy are independent of each other. In particular, we proved that, given the sampling times of the update packets, the Maximum Age First (MAF) scheduling strategy provides the best age performance among all scheduling strategies.

This transforms our overall optimization problem into an optimal sampling problem, given that the decision policy follows the MAF scheduling strategy. Interestingly, we showed that the zero-wait sampler

(recall that this is where a packet is generated once the channel is idle) is optimal for minimizing the TaPA, while it does not always minimize the TaA. We utilized Dynamic Programming (DP) to investigate the optimal sampler for minimizing the TaA. Finally, we provided in [7] an approximate analysis of Bellman’s equation to approximate the TaA-optimal sampler by a water-filling solution and demonstrate its efficacy through numerical evaluations.

6. MULTI-HOP WIRELESS NETWORKS

In most works studying AoI including ours, the goal has been to derive the age averaged over time for various single queue models. However, given how most communications these days happen over multiple hops, we wanted to extend our thinking and focus on analyzing the average AoI for queues in tandem emulating multi-hop paths. We wanted to gain insights into the performance of real-time status updating systems operating over a multi-hop network.

Fig. 16 shows a source that generates update packets as a rate λ Poisson process and sends them to the first queue. The packets flow through a series of queues serving packets at rate μ_i until they reach the monitor, as depicted. There is no preemption in this system, such that a packet in service completes service before the server becomes available for another packet, and a packet that arrives to a full queue is dropped. Aside from this queue overflow, there are no other packet losses under this model.

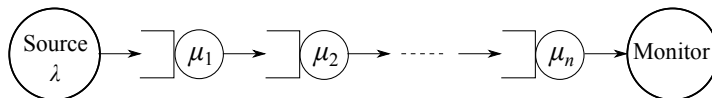


Fig. 16 – Queues in tandem.

With this backdrop, we started by deriving the average age for two and three $M/1/1$ non-preemptive queues in tandem, with the notation indicating that each of the queues have a capacity equal to 1 and there are no exogenous arrivals to the queues further down in tandem. This is done using a recently developed approach called Stochastic Hybrid Systems (SHS). The SHS method has the potential to lead to a generalized methodology for computing AoI for various queueing scenarios. All that is needed is a finite-state Markov chain, which can be used to generate a system of linear differential equations that describe the temporal convergence of the higher order moments and moment-generating functions of the AoI process. Limiting the queue capacity is based on the intuition that allowing updates to age in a queue may not be efficient when the objective is to maximize information freshness at the destination, since the queues will not store obsolete packets unnecessarily. This has the additional advantage of making the analysis of the average age more tractable. In addition, we believe this should approximate the performance of a non-preemptive Last Generated First Served (LGFS), and likely outperform it. To verify our theoretical result, we simulated the single capacity non-preemptive queues in tandem. Results in Fig. 17 show that the theory agrees with simulation quite well.

Next, we considered the case of infinite capacity first-come, first served queues in tandem, with Poisson arrival process and exponential servers, described as $M/1/\infty$ in [8], with the same caveat that there are no exogenous arrivals down the line. Unfortunately, applying the SHS approach in this case made the problem intractable, as this resulted in an explosion of states in the SHS model. Therefore, we went back to the traditional method of deriving the average age using a graphical argument, whereby we evaluated the correlations between system times and inter-arrival or inter-departure times. Again, in order to verify the theoretical result, we simulated these infinite capacity non-preemptive queues in tandem, as shown in Fig. 18(a).

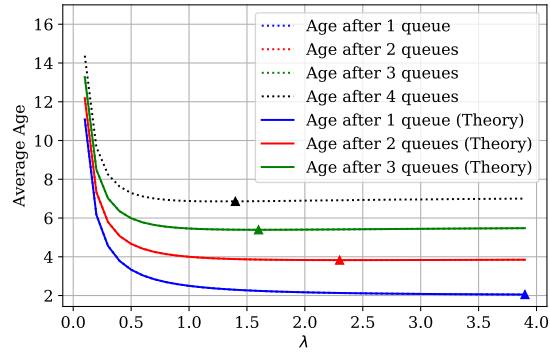


Fig. 17 – Average age vs. λ , $\mu = 1$, queue capacity of 1. Simulation minima indicated by Δ .

Finally, in order to develop a holistic understanding of this problem space, we studied the case of capacity- k non-preemptive queues in tandem. We generated simulation results for capacity- k queues in tandem as shown in Fig. 18(b), as we lack theoretical expressions of average age for arbitrary values of k .

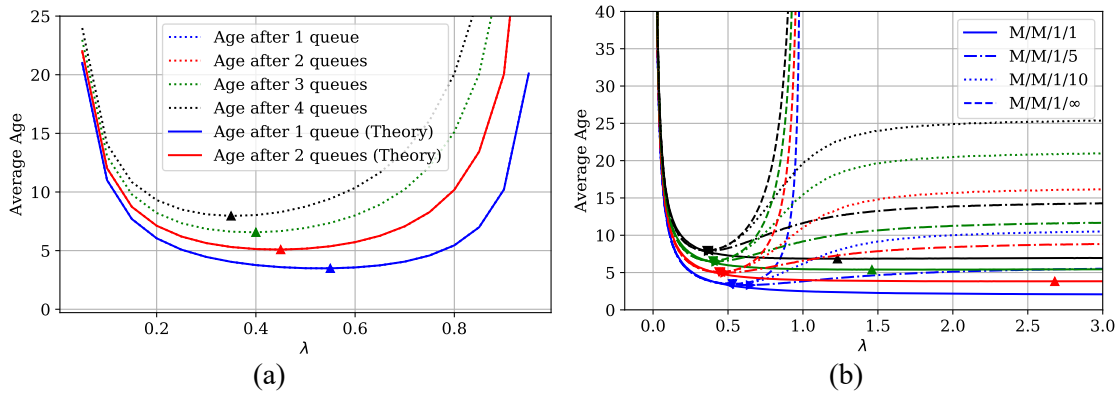


Fig. 18 – Average age vs. λ , $\mu = 1$: (a) infinite queue capacity, and (b) k queue capacity.

7. FUTURE RESEARCH DIRECTION

Our research team has been at the forefront of the emerging field of Age of Information (AoI), which measures the freshness of information in systems of a status updating type. The age metric is a breakthrough in characterizing and analyzing the performance of status monitoring applications, for which the traditional notion of delay is inadequate. While our work in this project has established a clear need for network control that is influenced by information freshness, such as in monitoring for situational awareness, there has been no clear approach yet towards effectively integrating the concepts from AoI into existing military networking architectures and scenarios, or how to prepare future architectures to best satisfy the need for freshness along with other mission-critical objectives. Additionally, it is important that the current state of the AoI research must be extended to the larger networking problem space. Specifically, in the future, we will look to extending the state of the Age of Information research to analyze the limits and tradeoffs within a heterogeneous networking environment, and incorporate various objectives, constraints, impairments, and unknowns. This integrated approach will enable the transition of these freshness concepts into existing and future military systems.

REFERENCES

1. C. Kam, S. Kompella, G. D. Nguyen, J. E. Wieselthier, A. Ephremides, “Towards an Effective Age of Information: Remote Estimation of a Markov Source,” in *Proc. IEEE INFOCOM Workshop on Age of Information*, pp.367–372, Apr. 2018.
2. C. Kam, S. Kompella, G.D. Nguyen, J.E. Wieselthier and A. Ephremides, “Towards an ‘Effective Age’ Concept,” in *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2018)*, Kalamata, Greece, June 2018.
3. C. Kam, S. Kompella, and A. Ephremides, “Learning to Sample a Signal through an Unknown System for Minimum AoI,” in *Proc. IEEE INFOCOM Workshop on Age of Information*, Paris, France, Apr. 2019.
4. C. Kam, S. Kompella, and A. Ephremides, “Age of Incorrect Information for Remote Estimation of a Binary Markov Source,” in *Proc. IEEE INFOCOM Workshop on Age of Information*, July 2020.
5. G.D. Nguyen, S. Kompella, C. Kam, J.E. Wieselthier and A. Ephremides, “Information Freshness over an Interference Channel: A Game Theoretic View,” in *Proc. IEEE INFOCOM*, Apr. 2018
6. G.D. Nguyen, S. Kompella, C. Kam, and J. E. Wieselthier, “Information Freshness over a Markov Channel: The Effect of Channel State Information,” in *Elsevier Ad-Hoc Networks Journal*, Vol. 86, pp. 63—71, 2019.
7. A.M. Bedewy, Y. Sun, S. Kompella, and N.B. Schroff, “Age-optimal Sampling and Transmission Scheduling in Multi-Source Systems,” in *Proc. ACM MobiHOC*, pp 1248—1257, 2019.
8. C. Kam, and S. Kompella, “On the Age of Information for Non-Preemptive Queues in Tandem,” in submission, *IEEE Internet of Things Journal*.