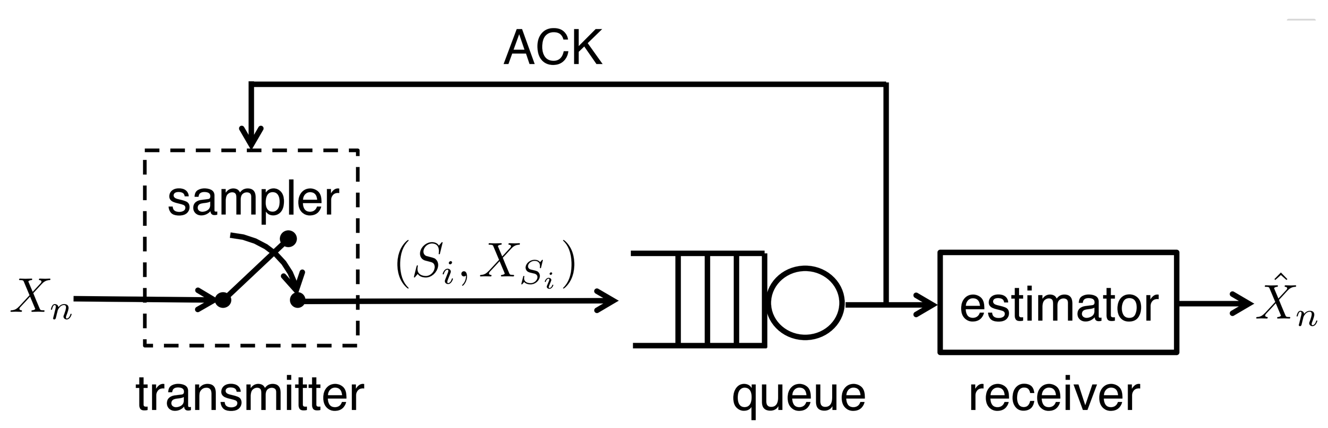
Information Aging through Queues: A Mutual Information Perspective

Yin Sun and Benjamin Cyr, Dept. ECE, Auburn University

Motivation

- Information is often most valuable when it is fresh.
  - E.g., car location for autonomous driving, interest-rate movements for stock trading

System Model



- Samples of discrete-time process (X_n) sent to receiver.
- Channel modeled as a FIFO queue with discrete i.i.d. service times (Y_i)
- Sample i generated at time S_i and delivered at time D_i

Definition: Age of Information (AoI)

$$\Delta_n = n - \max\{S_i : D_i \le n\}$$

Current time (n) — Generation time of freshness received sample ($\max\{S_i : D_i \le n\}$)

AoI is a **time difference** between transmitter and receiver

Issues:

- Signals vary over time with different speeds.
- **Time difference** in AoI is irrelevant of varying speed.

Our Contributions

- Define an information freshness metric based on *information structure* and *causality*
- Develop *optimal* online sampling policy to maximize the freshness of information of Markov sources
- Optimal sampling policy has a *nice structure*

Mutual Information as an Information freshness metric

- Samples received by time n:

$$\mathbf{W}_n = \{X_{S_i} : D_i \le n\}$$

- Information freshness metric:

$$I(X_n;\mathbf{W}_n) = H(X_n) - H(X_n|\mathbf{W}_n)$$

Mutual information between current signal value X_n and received samples $\mathbf{W}_n$.

- Interpretations:

1. **Measure freshness based on information, not time**
   - Fresh: $I(X_n;\mathbf{W}_n) \approx H(X_n)$
   - Obsolete: $I(X_n;\mathbf{W}_n) \approx 0$
2. **Reduction of binary experiments for inferring X_n**
   - Shannon code length without $\mathbf{W}_n \approx H(X_n)$
   - Shannon code length with $\mathbf{W}_n \approx H(X_n|\mathbf{W}_n)$
3. **Easy to compute for Markov sources X_n**

Lemma: If X_n is a time-homogenous Markov chain, then $I(X_n;\mathbf{W}_n) = I(X_n;X_{n-\Delta_n})$ is a non-negative, **non-increasing** function of the age $r(\Delta_n)$

- **Information "aging" = mutual information reducing**
- E.g., Gaussian Markov source: $X_n = aX_{n-1} + Z_n$

$$I(X_n;\mathbf{W}_n) = -\frac{1}{2}\log_2\left(1 - a^{2\Delta_n}\right)$$

- Binary Markov source: $X_n = X_{n-1} \oplus V_n$, $V_n \sim \text{Bernoulli}(q)$

$$I(X_n;\mathbf{W}_n) = 1 - h\left(\frac{1-(1-2q)^{\Delta_n}}{2}\right)$$

Online Sampling Problem

$$\bar{I}_{\text{opt}} = \sup_{\pi\in\Pi} \liminf_{N\to\infty} \frac{1}{N}\mathbb{E}\left[\sum_{n=1}^{N} I(X_n;\mathbf{W}_n)\right]$$

- Π : set of causal sampling policies
- Maximize the freshness of information over time
- Different from maximizing the rate of information in classic information theory

Optimal Sampling Policy of Markov sources

Theorem: If the service times Y_i are i.i.d., then there exists a threshold β such that the sampling policy

$$S_{i+1} = \min\{n\in\mathbb{N} : n \ge D_i, \mathbb{E}_{Y_{i+1}}[I(X_{n+Y_{i+1}};X_{S_i}|Y_{i+1}=y_{i+1})] \le \beta\}$$
$$= \min\{n\in\mathbb{N} : n \ge D_i, I(X_{n+Y_{i+1}};X_{S_i}|Y_{i+1}) \le \beta\}$$

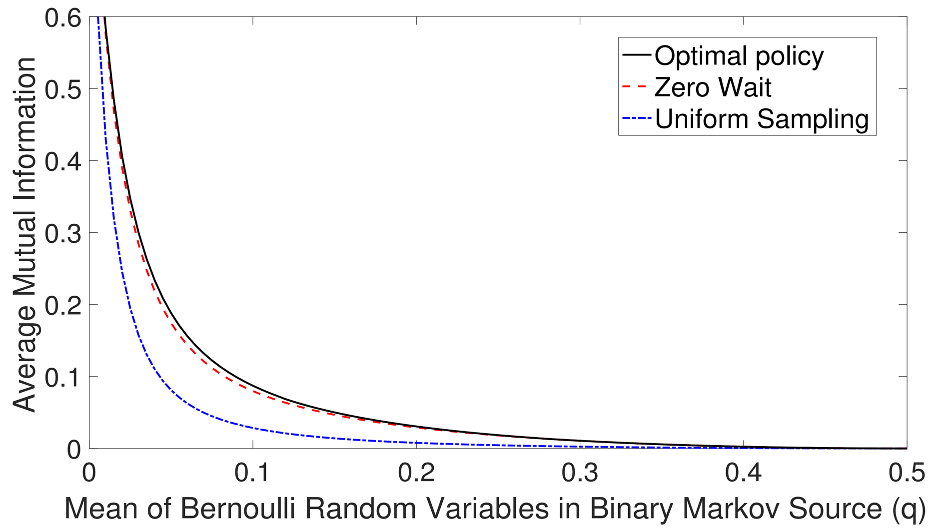
is optimal, where β is determined by solving

$$\beta = \frac{\mathbb{E}\left[\sum_{n=D_i}^{D_{i+1}-1} I(X_n;X_{S_i})\right]}{\mathbb{E}[D_{i+1}-D_i]}.$$

Further, β is exactly equal to the optimal value $\bar{I}_{\text{opt}}$.

- **Nice structure** of optimal sampler:
  - Must wait for previous sample to be delivered
  - Conditional mutual information reduces over time.
  - Must wait for the conditional mutual information drops below a threshold β
  - Threshold β equals to optimal objective value $\bar{I}_{\text{opt}}$
- Proof methodology:
  - Simpler proof, cleaner solution than [1].
  - Exact optimal for discrete-time problem, no approx.

Numerical Results



- Optimal policy much better than uniform sampling

References: [1]. Yin Sun, Elif Uysal-Biyikoglu, Roy D. Yates, C. Emre Koksal, and Ness B. Shroff, "Update or Wait: How to Keep Your Data Fresh", *IEEE Transactions on Information Theory*, Nov. 2017.

Poster available on Yin Sun's website. Expecting for your feedback!