TUT-5: TASK-ORIENTED AND SEMANTIC-AWARE COMMUNICATIONS AND NETWORKING FOR 6G

Nikolaos Pappas, Linköping University, Sweden
Yansha Deng, King's College London, United Kingdom
Jun Zhang, HKUST

LINKÖPING
UNIVERSITY

Outline

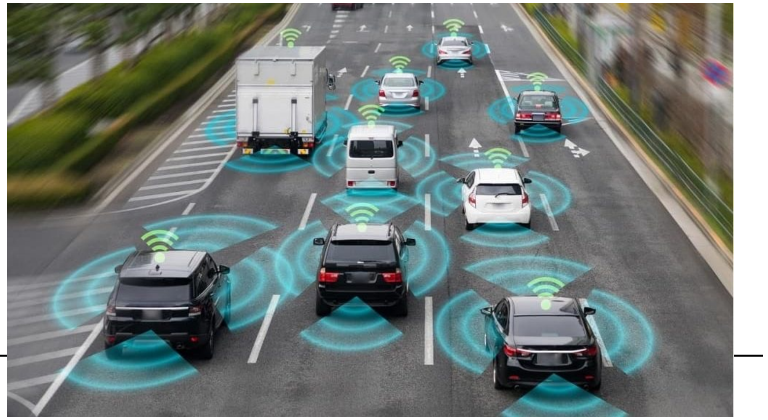
- Part I: Semantics of Information and Goal-oriented Communications (by Nikos)
- Part II: Task-Oriented Design for Metaverse (by Yansha)
- Part III: Task-oriented and Semantic-aware Communication for Edge Video Analytics (by Jun)

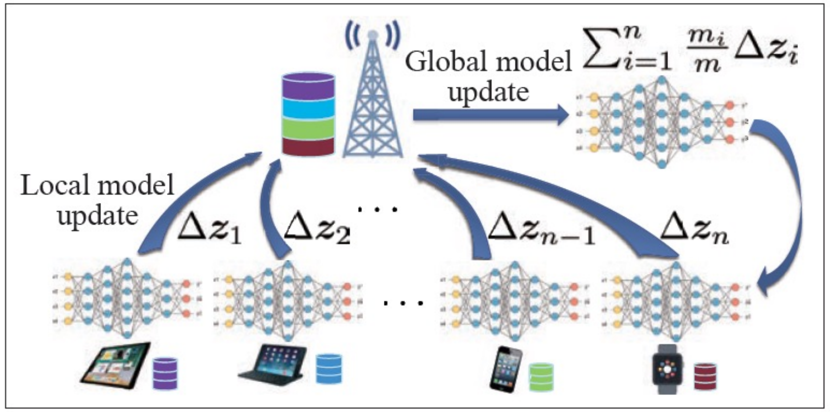
Part I - Outline

- Introduction and motivation
- Current approaches for measuring importance
  - Age of Information (AoI)
  - Non-linear AoI and Value of Information
  - Beyond AoI
- Ongoing research
  - Age of Actuation
  - Goal-Oriented Communication for Real-Time Tracking in Autonomous Systems
  - Fault detection in IoT
- Concluding remarks

Emerging wireless ecosystem

- Networked Intelligent Systems
  - Real-time autonomous systems
    - Sensor fusion, on time status updates, real time information reconstruction, network and device computation, traffic flows with synced requirements, autonomous interactions
- Distributed ML over wireless
  - Exchange of large datasets in a timely manner





Emerging wireless ecosystem

The envisioned use cases will stress future networks to deliver an unprecedented number of highly demanding requirements.

Emerging wireless ecosystem

The envisioned use cases will stress future networks to deliver an unprecedented number of highly demanding requirements.

- All wireless systems are built upon *fundamental principles* of reliable communications over noisy channels.
- In *existing communication paradigms*, the *main objective* is to *optimize* performance metrics, such as *throughput, delay, or packet l*oss.
- Quality of service (QoS) is provisioned through network over-provisioning and resource reservation control.

Emerging wireless ecosystem

The envisioned use cases will stress future networks to deliver an unprecedented number of highly demanding requirements.

- All wireless systems are built upon *fundamental principles* of reliable communications over noisy channels.
- In *existing communication paradigms*, the *main objective* is to *optimize* performance metrics, such as *throughput, delay, or packet loss*.
- Quality of service (QoS) is provisioned through network over-provisioning and resource reservation control.

- *Is not only about understanding the throughput-reliability-delay tradeoff.*

- *Maximizing throughput, minimizing delay are not enough for optimal operation in applications based on timely status updates, remote computations, and/or real-time event detection, etc.*

Towards Goal-oriented Semantic Communication

- Communication is about achieving specific goals.
- *Semantics: the semantic value of information is its usefulness in attaining a certain goal (pragmatics).*
- *Influence the relevance and effectiveness of the information we generate and communicate, depending on the objectives of the applications.*
- Utilize ***innate*** and ***contextual*** attributes of information.
- *A holistic redesign of the entire process of information generation, processing, transmission, and reconstruction.*

- [Chapter 7] A. Kosta, **N. Pappas**, V. Angelakis, "Age of Information: A New Concept, Metric, and Tool", Foundations and Trends in Networking: Vol. 12, No. 3, 2017.
- P. Popovski, O. Simeone, F. Boccardi, D. Gunduz, and O. Sahin, "Semantic-effectiveness filtering and control for post-5G wireless connectivity," *Journal of the Indian Institute of Science, 2020.*
- M. Kountouris, **N. Pappas**, "Semantics-Empowered Communication for Networked Intelligent Systems", *IEEE Communications Magazine, June 2021.*
- **N. Pappas**, M. Kountouris, "Goal-Oriented Communication for Real-Time Tracking in Autonomous Systems", *IEEE International Conference on Autonomous Systems (ICAS), Aug. 2021.*
- P. Popovski, F. Chiariotti, K. Huang, A. Kalor, M. Kountouris, **N. Pappas**, B. Soret, *"A Perspective on Time toward Wireless 6G", Proceedings of the IEEE, Aug. 2022.*

How to quantify importance of information

- **Real-time / time sensitive systems**: *Information usually has the highest value when it is fresh! (e.g., autonomous driving: info about location/speed/sensors)*

How to quantify importance of information

- **Real-time / time sensitive systems**: *Information usually has the highest value when it is fresh! (e.g., autonomous driving: info about location/speed/sensors)*
- *Age of Information (AoI):*
    - *AoI and its variants: simple, quantitative proxy metrics of information semantics*
    - ***Instrumental in establishing suboptimality of separate handling of sampling and communication***

How to quantify importance of information

- **Real-time / time sensitive systems**: *Information usually has the highest value when it is fresh! (e.g., autonomous driving: info about location/speed/sensors)*
- *Age of Information (AoI):*
  - *AoI and its variants: simple, quantitative proxy metrics of information semantics*
  - ***Instrumental in establishing suboptimality of separate handling of sampling and communication***
- Other cases such as
  - Quality of Information (QoI)
  - Value of Information (VoI)

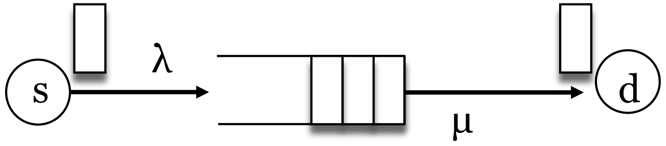
Definition and Modeling of AoI

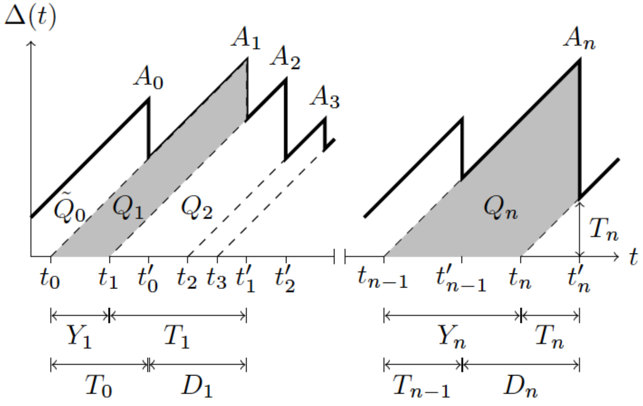
Definition of Age of Information (AoI)

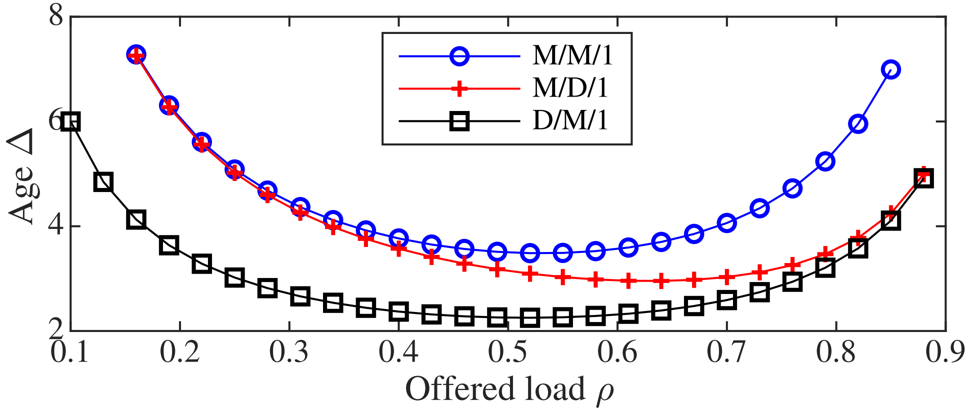
- AoI is an end-to-end metric that can characterize latency in status updating systems and applications and captures the timeliness of the information.
- An update packet with timestamp u has age $t-u$ at a time t.
- An update is fresh if its age is zero.
- When the monitor's freshest received update at time t has timestamp $u(t)$, the age is the random process $\Delta(t) = t - u(t)$.

Definition of Age of Information (AoI)

- AoI is an end-to-end metric that can characterize latency in status updating systems and applications and captures the timeliness of the information.
- An update packet with timestamp *u* has age *t*-*u* at a time *t*.
- An update is fresh if its age is zero.
- When the monitor's freshest received update at time *t* has timestamp *u(t)*, the age is the random process *Δ(t) = t - u(t)*.







Single-source and single-server systems – Packet management

- The arrival rate can be optimized to balance frequency of updates against congestion.
  - Departure from the external arrivals assumption.
- Study of lossy queues that may discard an arriving update while the server was busy or replace an older waiting update with a fresher arrival.
- ***Packet management inherently prioritizes some packets over others which is a first indication of different value of the packets thus the prioritization!***

- S. Kaul, R. Yates, M. Gruteser, “Status updates through queues”, CISS 2012.
- **N. Pappas**, J. Gunnarsson, L. Kratz, M. Kountouris, V. Angelakis, "Age of Information of Multiple Sources with Queue Management", IEEE ICC 2015.
- M. Costa, M. Codreanu, A. Ephremides, “On the age of information in status update systems with packet management”, IEEE Trans. Info. Theory 2016.
- A. Kosta, **N. Pappas**, A. Ephremides, V. Angelakis, "Age of Information Performance of Multiaccess Strategies with Packet Management", IEEE/KICS JCN, June 2019.

Non-linear Ageing

- AoI grows over time linearly
  - the performance degradation caused by information aging may not be a linear function of time.
- One way to capture the nonlinear behavior of information aging is to *define freshness and staleness as nonlinear functions of AoI*.
- A penalty function of the AoI is non-decreasing. *Outdated data is usually less desirable than fresh data.*

- Y. Sun, E. Uysal-Biyikoglu, R. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh", IEEE Trans. Inf. Theory, 2017.
- A. Kosta, **N. Pappas**, A. Ephremides, and V. Angelakis, "Age and value of information: Non-linear age case", IEEE ISIT 2017.
- Y. Sun and B. Cyr, "Sampling for data freshness optimization: Nonlinear age functions", IEEE/KICS JCN 2019.
- A. Kosta, **N. Pappas**, A. Ephremides, and V. Angelakis, "The cost of delay in status updates and their value: Non-linear ageing", IEEE Trans. Comm., 2020.

Cost of Update Delay (CoUD)

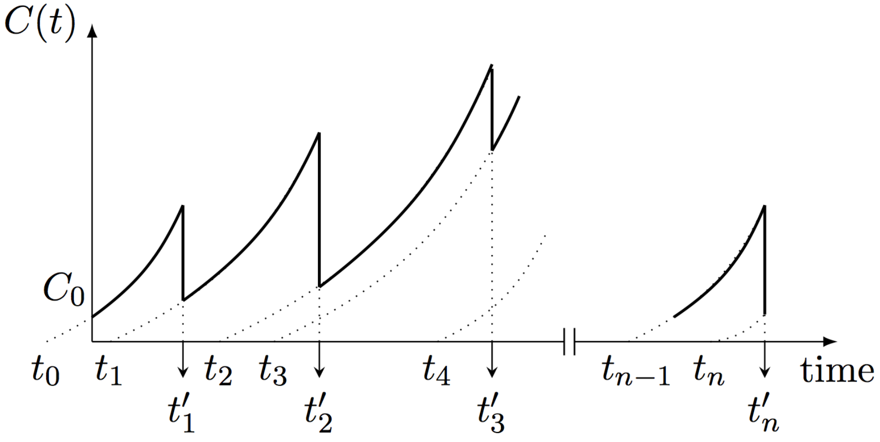
- CoUD metric associates the cost of staleness with the statistics of the source
- $C(t) = f_s(t-u(t))$
  - $f_s(t)$ is a monotonically increasing function
  - $u(t)$ timestamp of the most recently received update
- ***Different cost functions can capture different utilities***

- A. Kosta, **N. Pappas**, A. Ephremides, V. Angelakis, "Age and Value of Information: Non-linear Age Case", *IEEE ISIT* 2017.
- A. Kosta, **N. Pappas**, A. Ephremides, V. Angelakis, "The Cost of Delay in Status Updates and their Value: Non-linear Ageing", IEEE Trans. Comm., 2020.

Cost of Update Delay (CoUD):

Exponential case

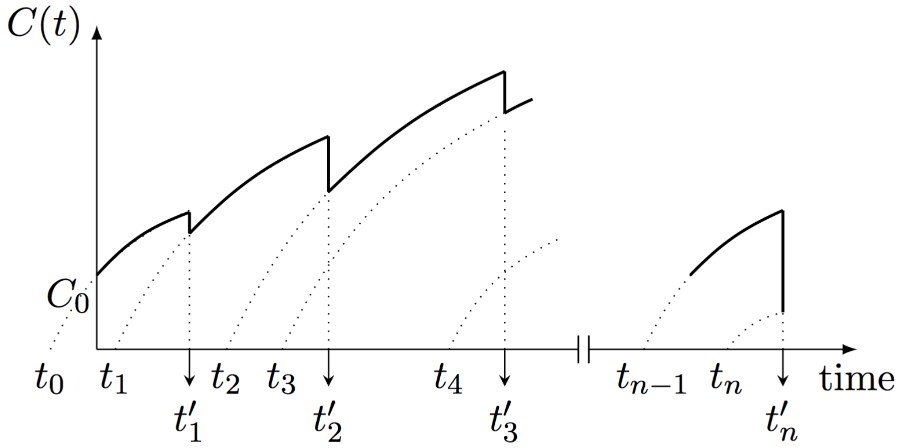
$f_s(t) = e^{\alpha t} - 1 \longleftrightarrow$ low autocorrelation



$$C_E = \mu(\rho - 1)\left(\frac{\alpha(\alpha - (\lambda + \mu))}{(\lambda - \alpha)(\alpha - \mu)^2} + \frac{1}{\alpha - \mu(1 - \rho)} + \frac{1}{\mu(1 - \rho)}\right)$$

Logarithmic case

$f_s(t) = \log(\alpha t + 1) \longleftrightarrow$ high autocorrelation



$$C_L = \frac{1}{\alpha(\lambda - \mu)^2}\left(e^{-\frac{\mu\rho}{\alpha}}\left(\mu(1 - \rho)Ei\left[-\frac{\mu}{a}\right](\alpha\mu + \lambda^2 - \lambda\mu)\,e^{\frac{\mu(\rho+1)}{\alpha}} - \alpha\mu^2(1 - \rho)Ei\left[-\frac{\lambda}{\alpha}\right]e^{\frac{\lambda+\mu\rho}{\alpha}}\right.\right.$$
$$\left.\left. - \alpha e^{\mu/\alpha}(\lambda - \mu)^2 Ei\left[-\frac{\mu(1 - \rho)}{\alpha}\right]\right) - \alpha\lambda(1 - \rho)(\mu - \lambda)\right)$$

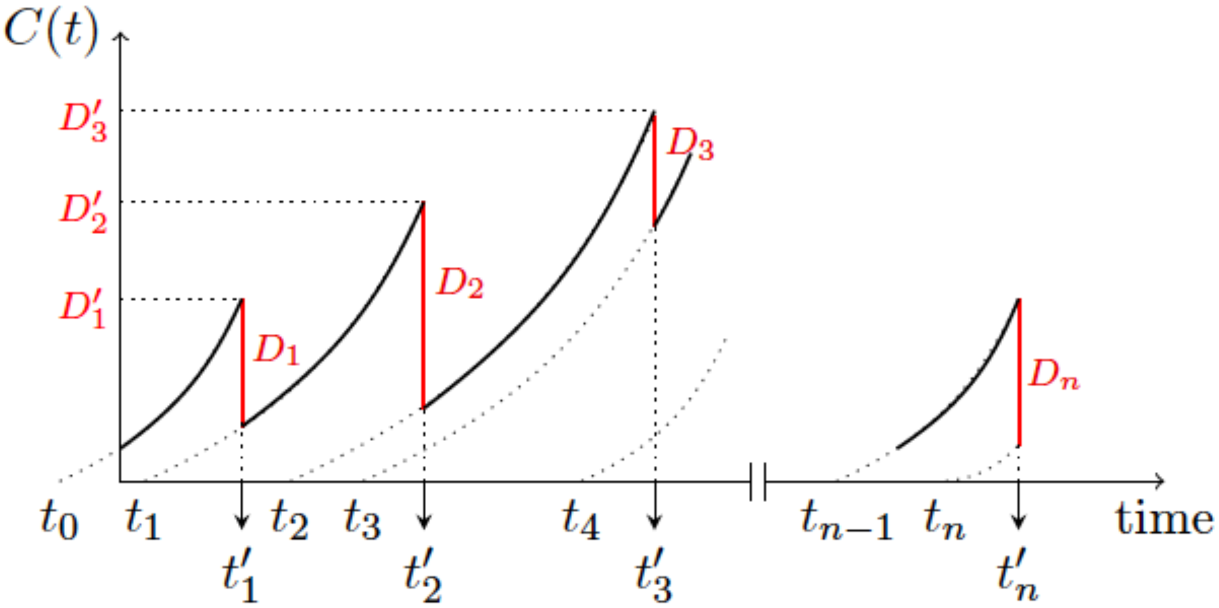
Value of Information of Update (VoIU)

- It captures *the degree of importance of an update*

$$V_i = \frac{f_s(t'_i - t_{i-1}) - f_s(t'_i - t_i)}{f_s(t'_i - t_{i-1})} = \frac{D_i}{D'_i}.$$

- In the linear CoUD case, VoIU is independent of the cost assigned per time unit →the Value is independent of the slope.

$$V_i = \lim_{t'_i \to t_i} \frac{f_s(t'_i - t_{i-1}) - f_s(t'_i - t_i)}{f_s(t'_i - t_{i-1})} = 1$$



- Linear case, the average VoIU for the M/M/1 system with an FCFS queue discipline.

$$V_P = \lambda \frac{(1-\rho)}{2\rho} {}_2F_1\left(1, 2; 3; 2 - \frac{1}{\rho}\right)$$

Pochhammer symbol

$${}_2F_1(a, b; c; z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!}$$

AoI and VoI in Control

- AoI considers only the timeliness!
- It has been shown that AoI alone does not capture the requirements of networked control loops.
- *Introduction of non-linear AoI facilitated the adoption in networked-control systems (NCS).*
- *VoI can reduce the estimation error in an NCS setup!*

O. Ayan, M. Vilgelm, M. Klügel, S. Hirche, and W. Kellerer, "Age-of-Information vs. Value-of-Information Scheduling for Cellular Networked Control Systems", 10th ACM/IEEE ICCPS 2019.

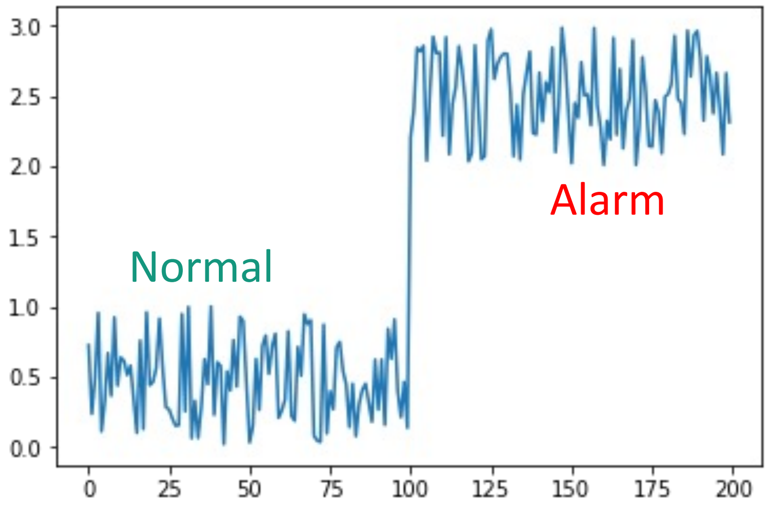
Extending AoI

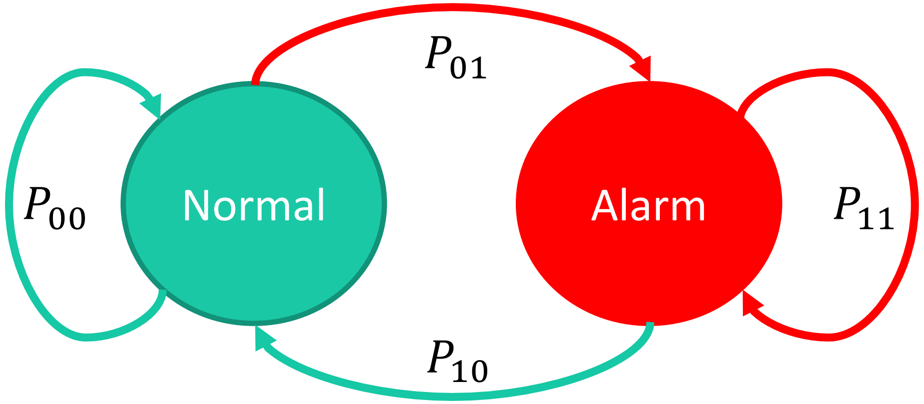
- The classical AoI does not capture properties of the source
  - Except timeliness itself which is a semantic property
- With non-linear AoI and the VoI, we go a step further
- Here we will discuss another extension of AoI

G. Stamatakis, **N. Pappas**, A. Traganitis, "Control of Status Updates for Energy Harvesting Devices that Monitor Processes with Alarms", IEEE GLOBECOM Workshops, Dec. 2019.

Stochastic process with alarms (or a two-state)

- Short timescale: Stochastic process Z_k evolves over discrete time k.
- Long timescale: A two state Markov Process
  - States: Normal, Alarm,
  - Geometrically distributed sojourn times in each state.
- The Alarm state indicates the need for more frequent updates
  - Closely follow/track the evolution of Z_k to make informed decisions.
- Examples of Z_k
  - the network load under normal operation and under a DoS attack.
  - Physical phenomena such as temperature, water levels, and air pollution.
  - *Tracking of a process in general.*
- Objective: Optimize the freshness of status updates at the destination while considering the energy resources. currently available as well as future demands for energy (especially during alarm periods)



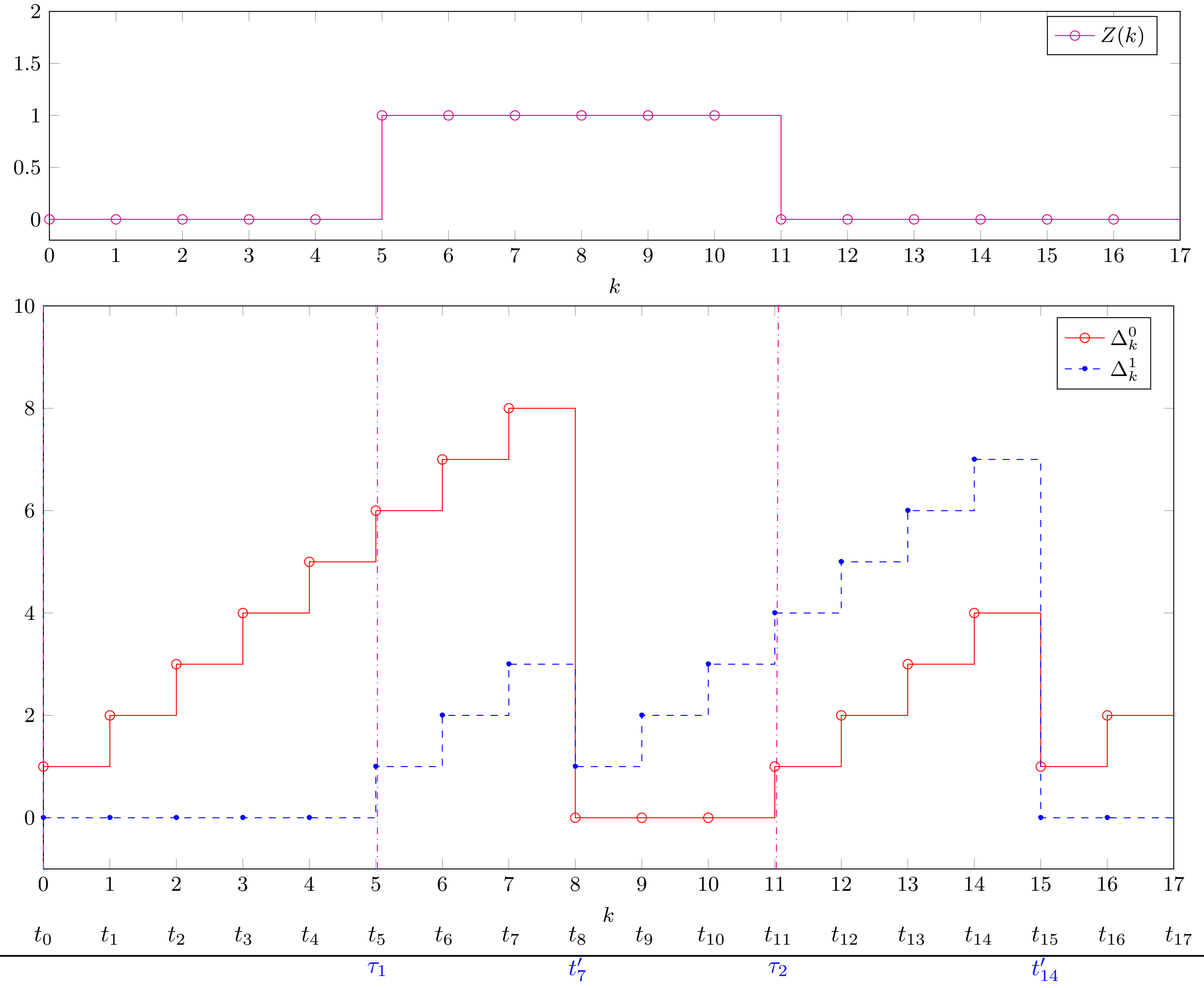


G. Stamatakis, **N. Pappas**, A. Traganitis, "Control of Status Updates for Energy Harvesting Devices that Monitor Processes with Alarms", IEEE GLOBECOM Workshops, Dec. 2019.

AoI for stochastic processes with alarms

- Extend the definition of AoI
  - The amount of time that has elapsed since the generation the last status update that has been successfully received by the monitor (typical)
  - The amount of time that has elapsed since the last state change of the stochastic process for which the destination in uninformed (new)
- Use two AoI variables, one for each state of the process $\Delta_{\mathrm{k}}^{z}, z \in \{0,1\}$.
- The destination knows the stochastic process to be in state Z_k^d.
  - Not necessarily the actual state of the stochastic process indicated with Z_k.

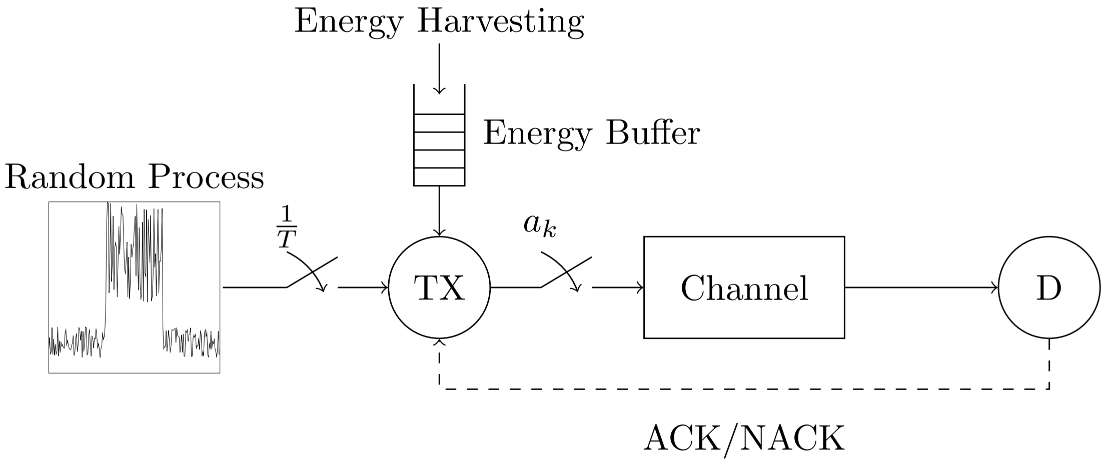
Extended AoI - illustration for the two-state process



$$\Delta_k^z = \begin{cases} k - U_k, & \text{if } z = Z_k^d \\ k - \tau_n, & \text{if } z \neq Z_k^d \text{ and } z = Z_k \\ 0, & \text{if } z \neq Z_k^d \text{ and } z \neq Z_k \end{cases}$$

System model

- At the beginning of the k-th timeslot the sensor samples/senses the stochastic process in order to assess Z_k.
- The sensor also considers
  - The state of the process known at the destination Z_k^d.
  - The energy stored at the energy buffer E_k
  - The values of both AoI variables $\Delta_\mathrm{k}^0, \Delta_\mathrm{k}^1$
- These features constitute the **state** of the system
  - $s_k = [Z_k, Z_k^d, E_k, \Delta_\mathrm{k}^0, \Delta_\mathrm{k}^1]$
- Given s_k the sensor must choose whether to transmit a fresh status update or not, $a_k \in \{0,1\}$.
- *Energy harvesting can be considered an abstraction to capture the availability of the transmitter.*



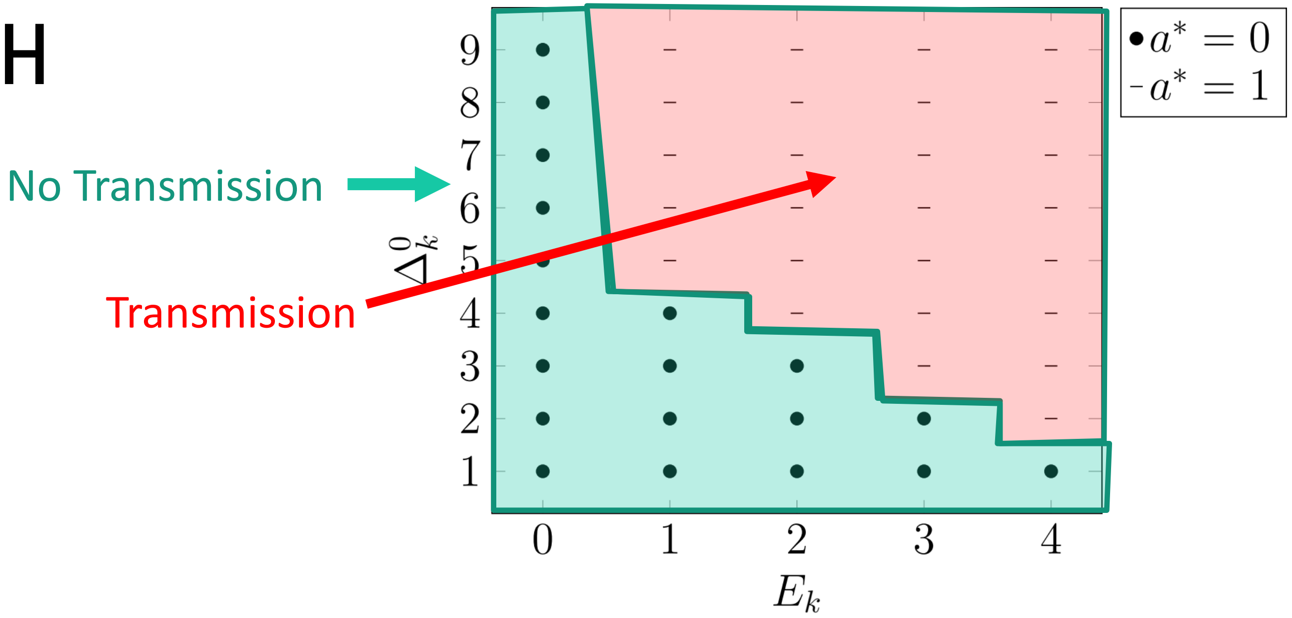
System model

- At the end of each time-slot a cost is paid by the sensor.
- The transition cost is an increasing function of Δ_k^0 and Δ_k^1.
- $g(\Delta_k^0, \Delta_k^1) = g_0(\Delta_k^0) + g_1(\Delta_k^1)$
  - **$g_1(\cdot)$ increases faster than $g_0(\cdot)$.**
  - ***This expresses the need for frequent status updates when in alarm state →The value of information in that case is higher!***
- Examples:
  - $g(\Delta_k^0, \Delta_k^1) = (1 - Z_k)\Delta_k^0 + Z_k\left(\Delta_k^1\right)^2$
  - i.e., cost is a function of the true state of the stochastic process and not the one perceived by the destination.
  - $g(\Delta_k^0, \Delta_k^1) = \Delta_k^0 + \left(\Delta_k^1\right)^2$
  - cost considers both AoI variables simultaneously (Upcoming work)

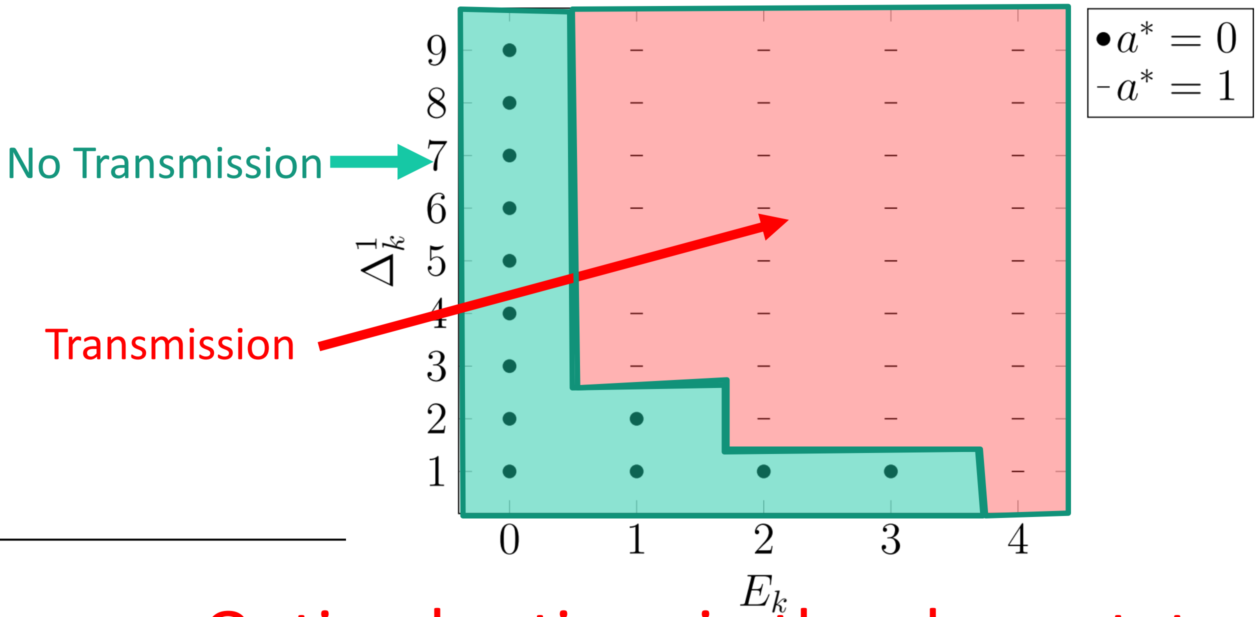
- Objective: Find an optimal policy that, given s_k, decides when to transmit a status update to minimize the discounted transition costs accumulated over an infinite horizon.
- The problem is a *Markov Decision Process*, and the optimal policy can be found via the Value Iteration algorithm.
- *Curse of dimensionality can be circumvented by utilizing structural results for the optimal policy.*

Optimal Policy - Low Probability EH ($P_e = 0.4$) - $P_s = 0.8$

- Scenario
  - the process spends most of its time in normal mode with relative short periods of alarm states.
    - State transition matrix $P_z = \begin{bmatrix}0.9 & 0.1\\0.2 & 0.8\end{bmatrix}$
- *The optimal policy will save energy in the normal state in order to be able to transmit in the alarm state*



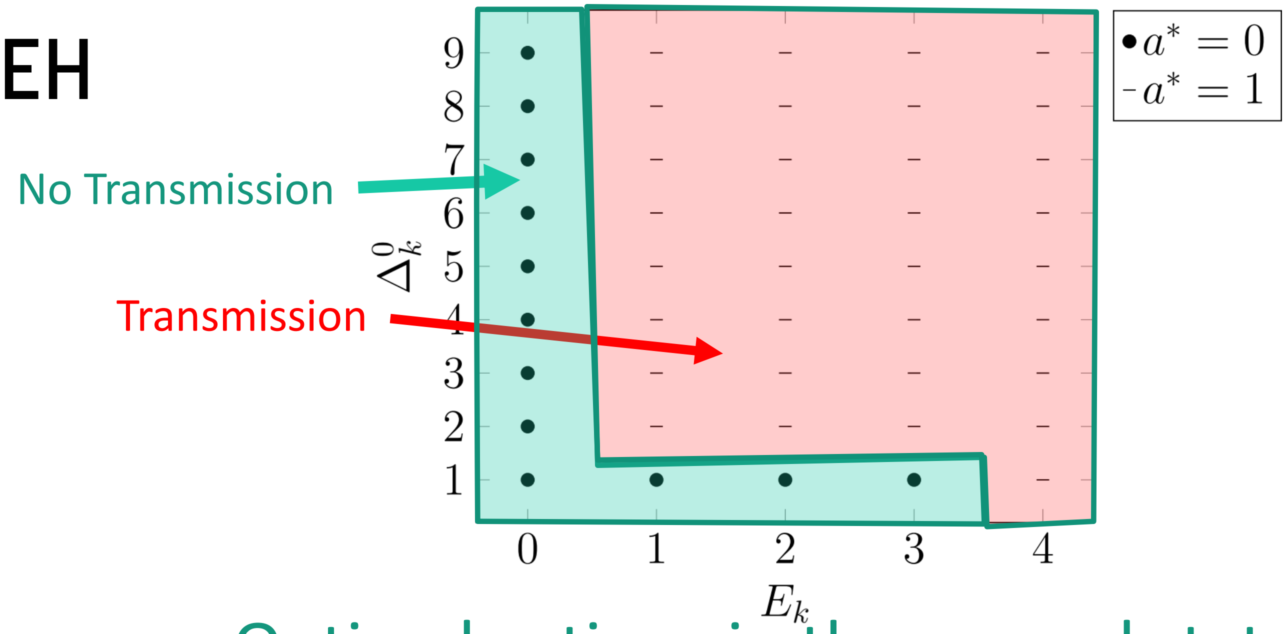
Optimal actions in the normal state



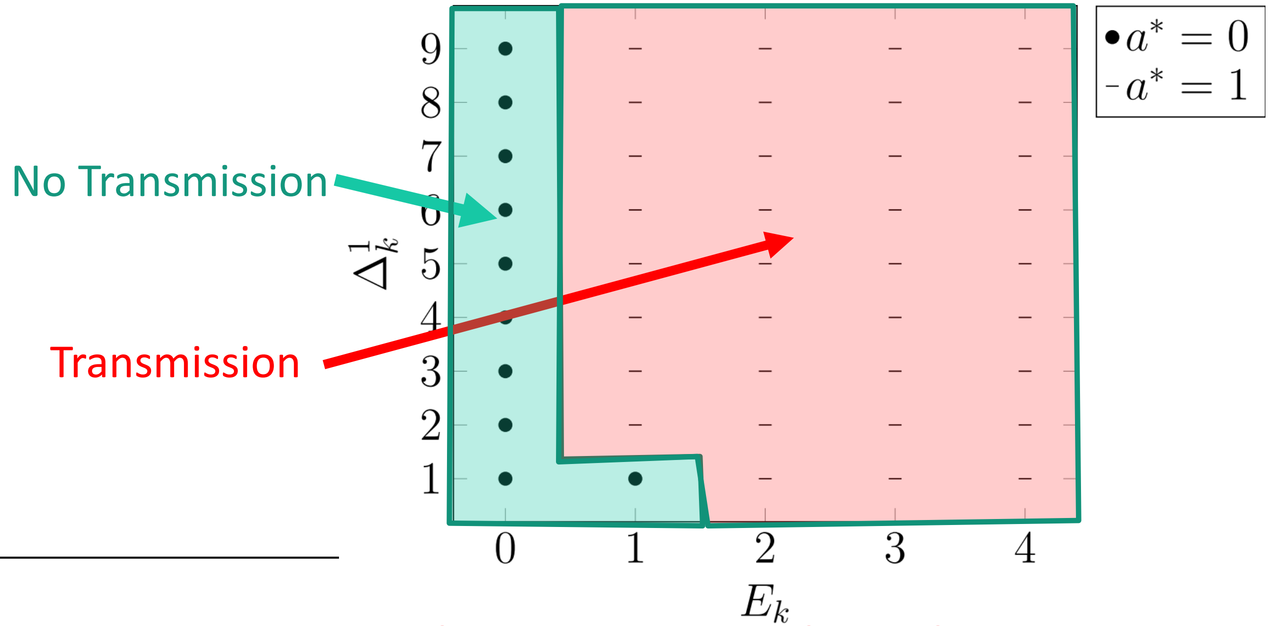
Optimal actions in the alarm state

Optimal Policy - High Probability EH ($P_e = 0.8$) - $P_s = 0.8$

- Scenario
  - the process spends most of its time in normal mode with relative short periods of alarm states.
    - State transition matrix $P_z = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$
- *Energy saving is less important when EH occurs with high probability!*



Optimal actions in the normal state



Optimal actions in the alarm state

Discussion

- *Some other metrics appeared after that work*
  - Age of Incorrect Information
  - Pull based AoI

- A. Maatouk, S. Kriouile, M. Assaad and A. Ephremides, "The Age of Incorrect Information: A New Performance Metric for Status Updates", IEEE/ACM Trans. on Networking 2020.
- J. Holm, A. E. Kalør, F. Chiariotti, B. Soret, S. Jensen, T. Pedersen, and P. Popovski, "Freshness on demand: Optimizing Age of Information for the query process", IEEE ICC 2021.
- F. Li, Y. Sang, Z. Liu, B. Li, H. Wu, and B. Ji, "Waiting but not aging: Optimizing information freshness under the pull model", IEEE/ACM Trans. on Networking 2021.

- *Another relevant metric is the* ***Version AoI***

- R. D. Yates, "The Age of Gossip in Networks," IEEE International Symposium on Information Theory (ISIT), 2021.
- B. Buyukates, M. Bastopcu and S. Ulukus, "Version Age of Information in Clustered Gossip Networks," IEEE Journal on Selected Areas in Information Theory 2022.
- E. Delfani and **N. Pappas**, "Version Age-Optimal Cached Status Updates in a Gossiping Network with Energy Harvesting Sensor". TechRxiv, 12-Dec-2022, doi: 10.36227/techrxiv.21696617.v1.

Remarks and future directions

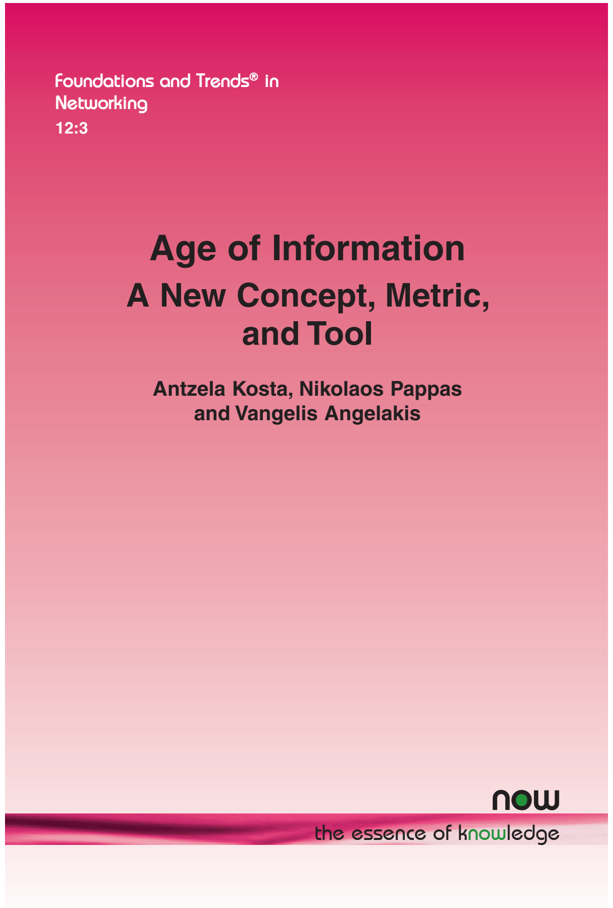
- AoI has emerged as an end-to-end performance metric for systems that employ status updates.

- Introduction of **information freshness** requirements ***will create systems that work smarter than harder***, so they will be **more *effective***.
  - The updating process should not underload nor overload the system.
  - The system should process new updates rather than old.
  - The system should avoid processing updates without sufficient **novelty**.

- A. Kosta, **N. Pappas**, V. Angelakis, "Age of Information: A New Concept, Metric, and Tool", Foundations and Trends in Networking: Vol. 12, No. 3, 2017.

Remarks and future directions

- There are still many interesting research directions
  - Definition of effective age (term coined by *Prof. Ephremides* in ITA 2015)
  - Sampling and remote reconstruction
  - Deploying of AoI in machine learning
  - Security
- It provides stronger connections with areas such as Signal Processing
- Metrics that can capture the requirements of Wireless Networked Control Systems
- AoI is one of the dimensions of *semantics-empowered communications*!
  - AoI is an ***innate*** attribute of information
  - Non-linear AoI is a ***contextual*** attribute
    - M. Kountouris, **N. Pappas**, "Semantics-Empowered Communication for Networked Intelligent Systems", *IEEE Communications Magazine, June 2021.*
    - P. Popovski, F. Chiariotti, K. Huang, A. Kalor, M. Kountouris, **N. Pappas**, B. Soret, *"A Perspective on Time toward Wireless 6G", Proceedings of the IEEE, Aug. 2022.*
    - ***N. Pappas**, M. A. Abd-Elmagid, B. Zhou, W. Saad, H. S. Dhillon, "Age of Information: Foundations and Applications", Cambridge University Press, Feb. 2023.*

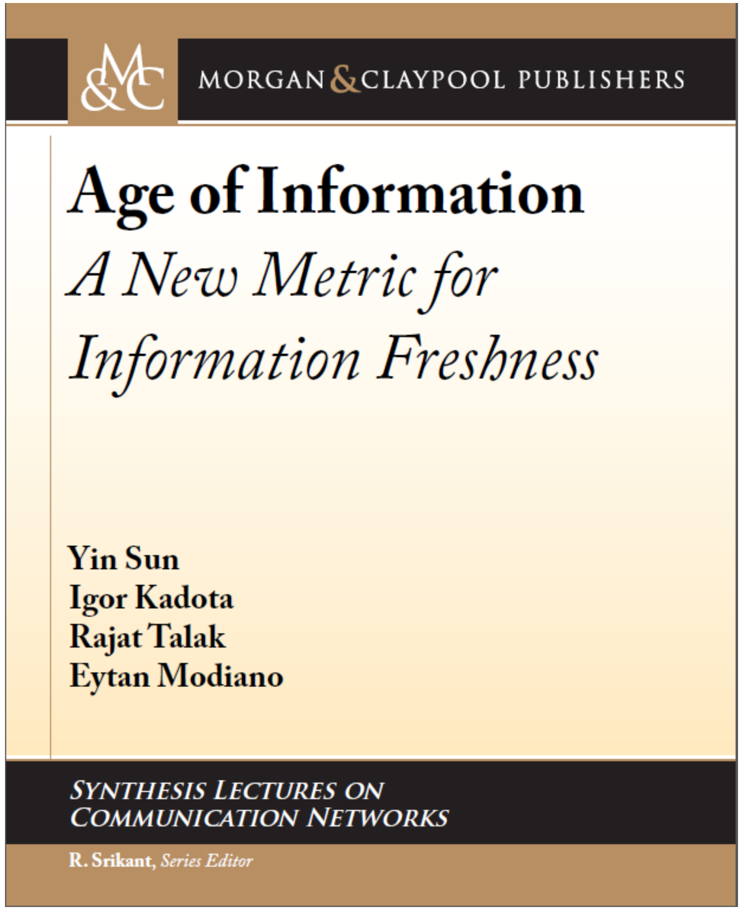


Foundations and Trends® in Networking
12:3

Age of Information
A New Concept, Metric, and Tool

Antzela Kosta, Nikolaos Pappas and Vangelis Angelakis

now
the essence of knowledge



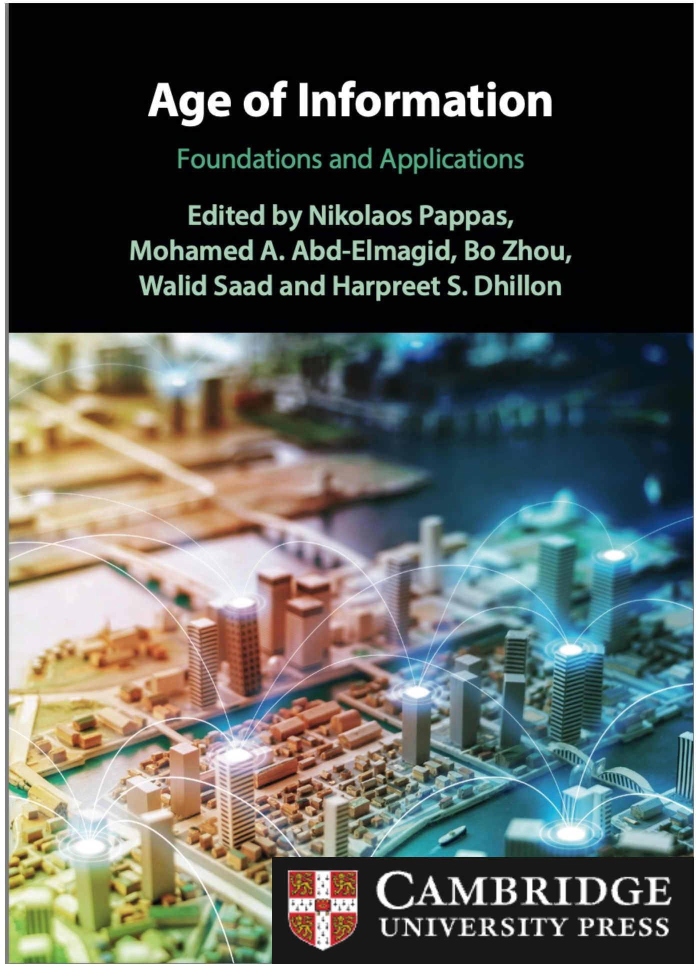
MORGAN & CLAYPOOL PUBLISHERS

Age of Information
A New Metric for Information Freshness

Yin Sun
Igor Kadota
Rajat Talak
Eytan Modiano

SYNTHESIS LECTURES ON COMMUNICATION NETWORKS

R. Srikant, *Series Editor*



Age of Information
Foundations and Applications

Edited by Nikolaos Pappas, Mohamed A. Abd-Elmagid, Bo Zhou, Walid Saad and Harpreet S. Dhillon

CAMBRIDGE UNIVERSITY PRESS

Semantics beyond Age of Information

- AoI is a proxy towards semantics communications
- AoI is a semantic metric
    - Captures timeliness of information (semantic property)
    - *However, the freshest information may not be the most useful.*

- Value of information
    - Non-linear AoI/VoI is a step towards that direction
    - Extended AoI capturing tracking of a process
- Accuracy and importance of information
- Joint aspects of timeliness and accuracy/importance

General description of *Semantics of Information (SoI)* metric

- A comprehensive system metric, *Semantics of Information (SoI)*, which captures the significance and usefulness of information w.r.t the goal of data exchange and the application requirements.
- Information attributes, which can be decomposed into *innate* (objective) and *contextual* (subjective).
- **Innate** are the attributes inherent to information regardless of its use, such as AoI, precision, correctness.
- **Contextual** are attributes that depend on the particular context or application for which information is being used.
  - For example, timeliness – as a function of AoI, accuracy (distortion), perception via divergence or distance functions.

N. Pappas, M. Kountouris, "Goal-Oriented Communication for Real-Time Tracking in Autonomous Systems", *IEEE International Conference on Autonomous Systems (ICAS), Aug. 2021.*

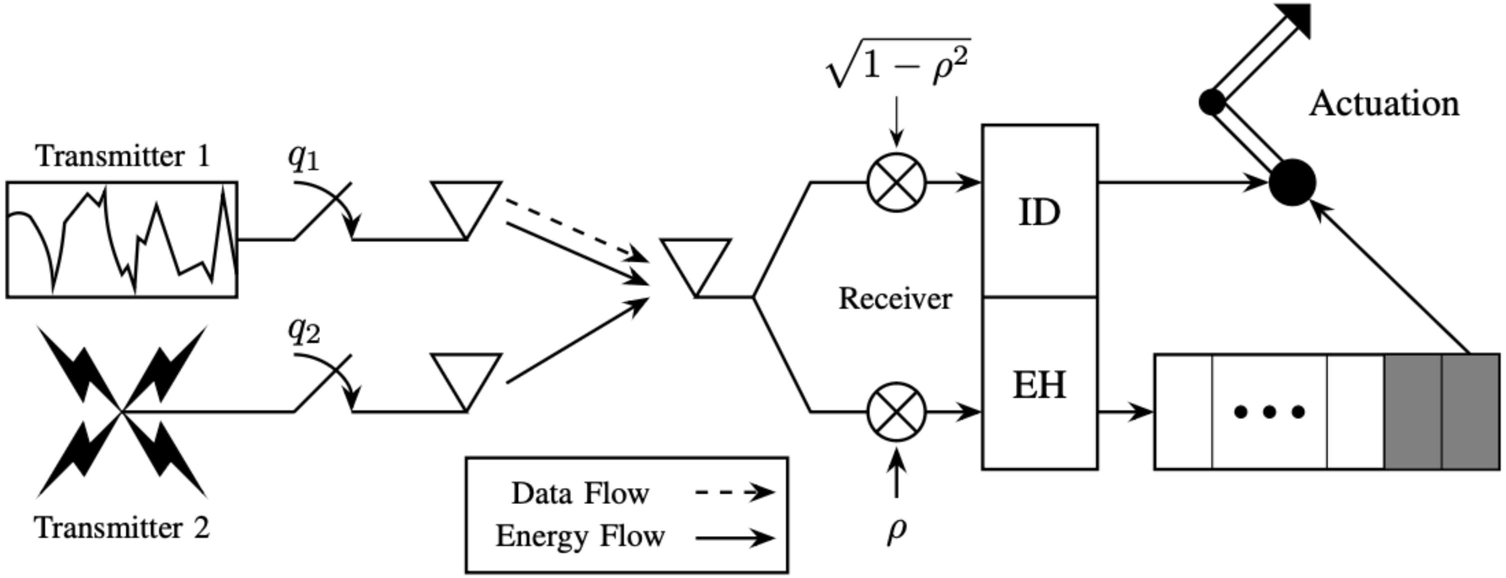
Age of Actuation in a Wireless Power Transfer System

A. Nikkhah, A. Ephremides, N. Pappas, ‘‘Age of Actuation in a Wireless Power Transfer System", IEEE INFOCOM - 6th Age of Information Workshop, May 2023.

LINKÖPING
UNIVERSITY

System Model

- Observations from an external process are transmitted through status updates to a battery-powered receiver.
- The receiver is informed about the status of the process and if there is sufficient energy, uses them to perform an actuation to achieve a goal.
- We consider a wireless power transfer model.
- We propose a new metric, the *Age of Actuation (AoA)* which is relevant when the receiver utilizes the status updates to perform actions in a timely manner.



Age of Actuation (AoA)

- In order to perform an action, we need to receive a status update and a non-empty battery at the receiver.

$$A(t) = t - a(t)$$

time of the last performed actuation

$$\bar{A} = \begin{cases} \bar{A}_1 = \frac{1}{P_{\mathcal{D}}} & \frac{P_{\bar{\mathcal{D}},\mathcal{E}}}{P_{\mathcal{D},\bar{\mathcal{E}}}} \geq 1 \\ \bar{A}_2 = \frac{1}{P_{\mathcal{D}} \frac{P_{\bar{\mathcal{D}},\mathcal{E}}}{P_{\mathcal{D},\bar{\mathcal{E}}}} + P_{\mathcal{D},\mathcal{E}} \left(1 - \frac{P_{\bar{\mathcal{D}},\mathcal{E}}}{P_{\mathcal{D},\bar{\mathcal{E}}}}\right)} & \frac{P_{\bar{\mathcal{D}},\mathcal{E}}}{P_{\mathcal{D},\bar{\mathcal{E}}}} < 1 \end{cases}$$

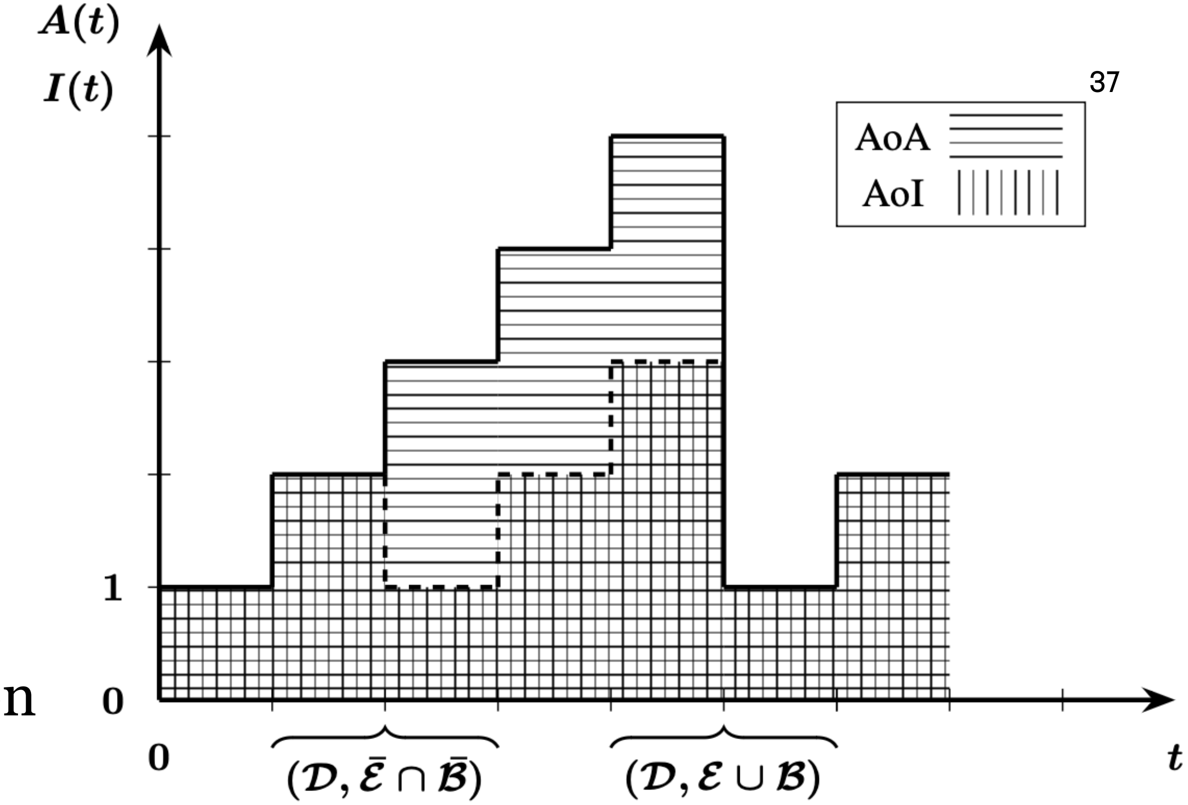


Fig. 2. The evolution of AoA and AoI metrics.

Numerical results

SUCCESS PROBABILITIES OF DATA AND ENERGY TRANSMISSION.

	P_{d1}	P_{d12}	P_{e2}	P_{e12}
Setup 1	1	0.62	0.20	0.23
Setup 2	1	0.34	0.60	0.63

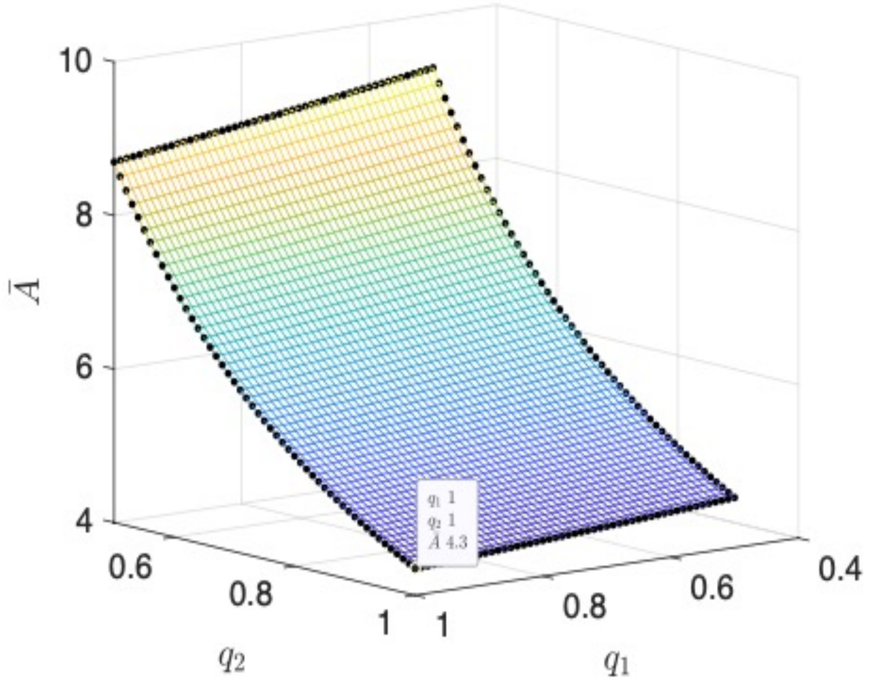


Fig. 3. Average AoA for the infinite-sized battery for the first setup. The minimum $\bar{A}^* = 4.3$ is achieved by $q_1^* = 1$ and $q_2^* = 1$.

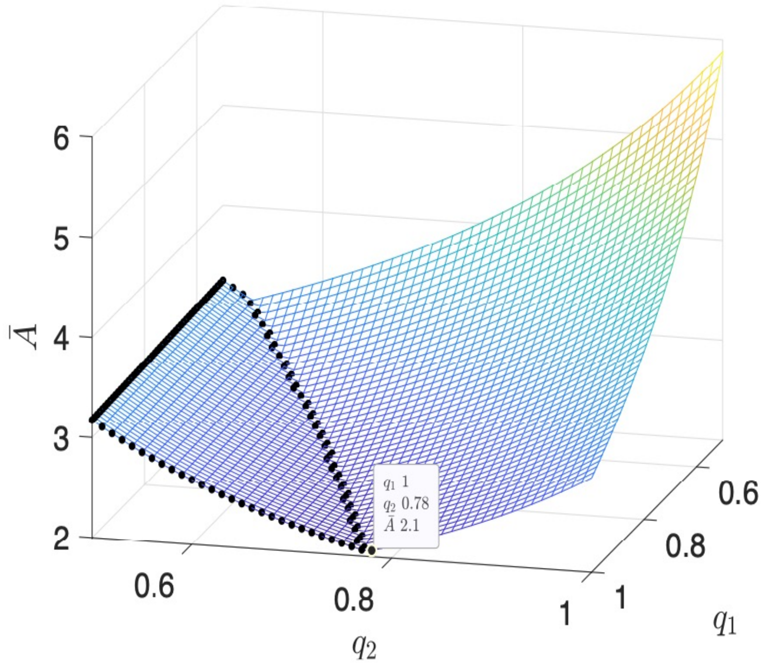


Fig. 4. Average AoA for the infinite-sized battery for the second setup. The minimum $\bar{A}^* = 2.1$ is achieved by $q_1^* = 1$ and $q_2^* = 0.78$.

Goal-Oriented Communication for Real-Time Tracking in Autonomous Systems

N. Pappas, M. Kountouris, ‘‘Goal-Oriented Communication for Real-Time Tracking in Autonomous Systems", IEEE International Conference on Autonomous Systems (ICAS), Aug. 2021.

LINKÖPING
UNIVERSITY

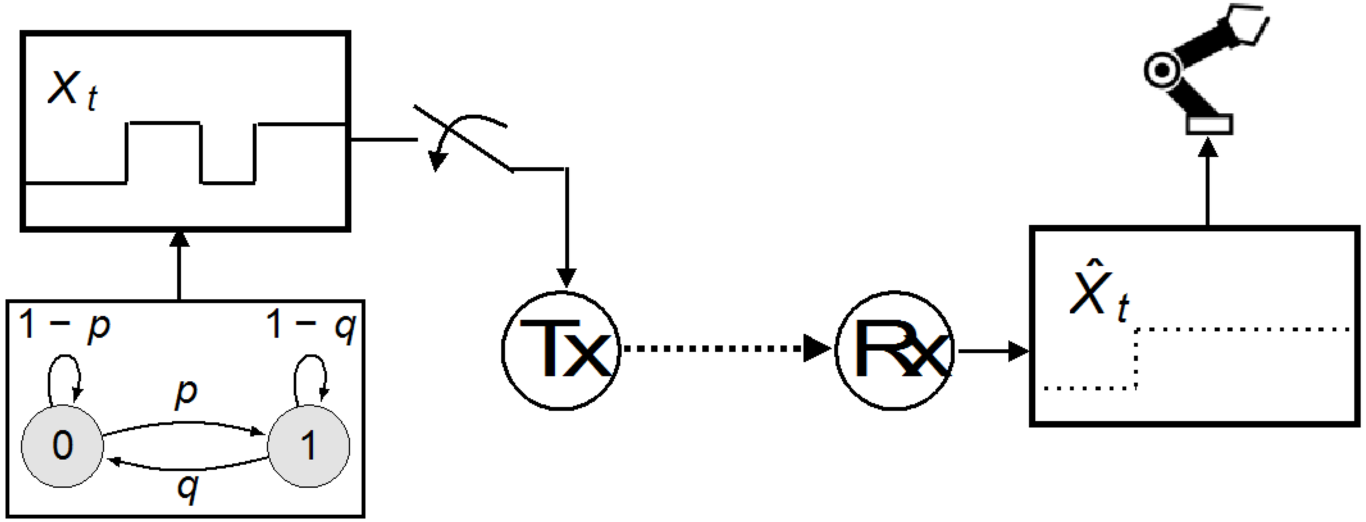
Introduction

- We consider real-time tracking and reconstruction of an information source.
- Real-time reconstruction is performed at the destination for remote actuation.
- Relevant setting for real-time applications in autonomous networked systems.

- We introduce *new goal-oriented, semantics-empowered sampling and communication policies*, which account for the temporal evolution of the source/process and the semantic and application-dependent value of data being generated and transmitted.

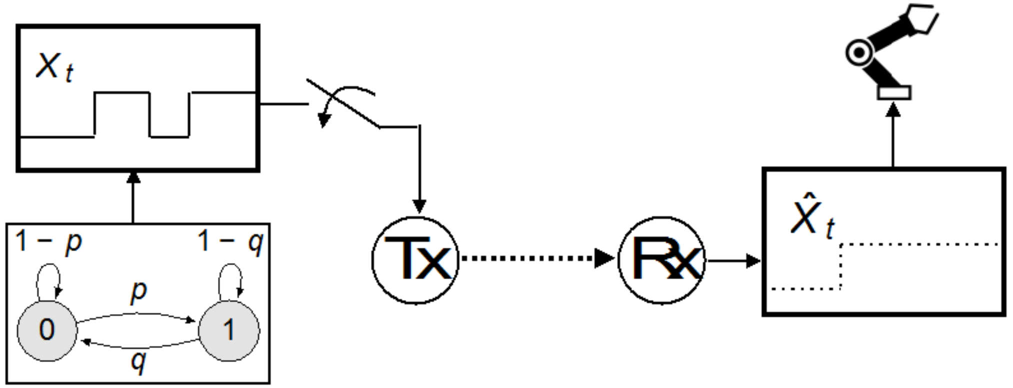
System Model

- A device monitors a two-state random process.
- The source initiates certain actions to the robotic object.
- The monitoring device samples and transmits status updates regarding the evolution of the source.
- *The application objective is to perform/maintain the actions of the original object in real-time.*



System Model

- Time is slotted.
- Wireless erasure communication channel with success probability:
  $p_s = \mathbb{P}\left(h_t = 1\right)$ where h_t is the channel realization.
- ACK/NACK instantaneous and error free.
- Information source, X_t, is modelled by a two state Markov Chain.
- X_t is reconstructed at the destination, $\hat{X}_t$, to perform actuation.
- The action of transmitting a sample is $\alpha_t^{\text{tx}} = 1$, otherwise, the transmitter remains silent $\alpha_t^{\text{tx}} = 0$.



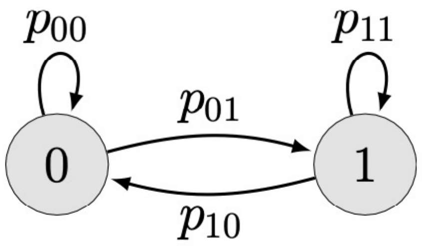
Key Performance Metrics

- **Real-time reconstruction error (innate)**: measures the discrepancy between the original and the reconstructed source in a timeslot

$$E_t = \mathbb{1}\left(X_t \neq \hat{X}_t\right) = \left|X_t - \hat{X}_t\right|$$

Time-averaged

$$\bar{E} = \lim_{T\to\infty} \frac{\sum_{t=1}^{T} E_t}{T} = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{1}\left(X_t \neq \hat{X}_t\right)$$



$$p_{ji} = \mathbb{P}\left(E_{t+1} = j | E_t = i\right)$$

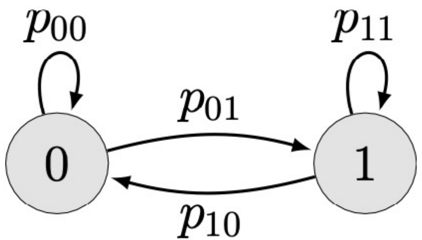
Key Performance Metrics

- **Real-time reconstruction error (innate)**: measures the discrepancy between the original and the reconstructed source in a timeslot

$$E_t = \mathbb{1}\left(X_t \neq \hat{X}_t\right) = \left|X_t - \hat{X}_t\right|$$

Time-averaged

$$\bar{E} = \lim_{T\to\infty} \frac{\sum_{t=1}^{T} E_t}{T} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{1}\left(X_t \neq \hat{X}_t\right)$$



$$p_{ji} = \mathbb{P}\left(E_{t+1} = j \middle| E_t = i\right)$$

- **Cost of actuation error (contextual)**: captures the significance of the error at the point of actuation. *Some errors may have larger impact than others.*
  - $C_{i,j}$ denotes the cost of being in state *i* at the original source and in *j* at the reconstructed, when E_t=*1*. In general $C_{0,1} \neq C_{1,0}$.

Average cost of actuation $\bar{C}_A = \pi_{(0,1)} C_{0,1} + \pi_{(1,0)} C_{1,0}$

Sampling and communication policies

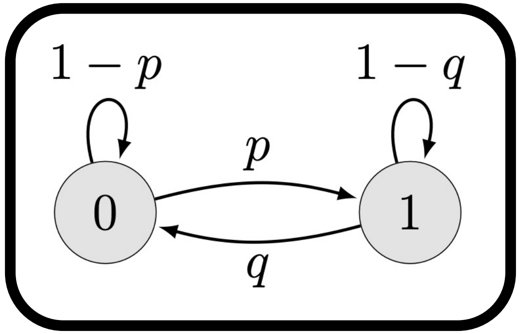
- Uniform: sampling is performed periodically, independently of the temporal evolution of the source.
  - *It is a process-agnostic policy that could result in missing several state transitions during the time interval between two collected samples.*
- Age-aware: the receiver triggers the acquisition and transmission of a new sample, once the AoI reaches a predefined threshold A_{th}.
  - *This policy is source-agnostic regarding the value of information but takes into account the timeliness.*

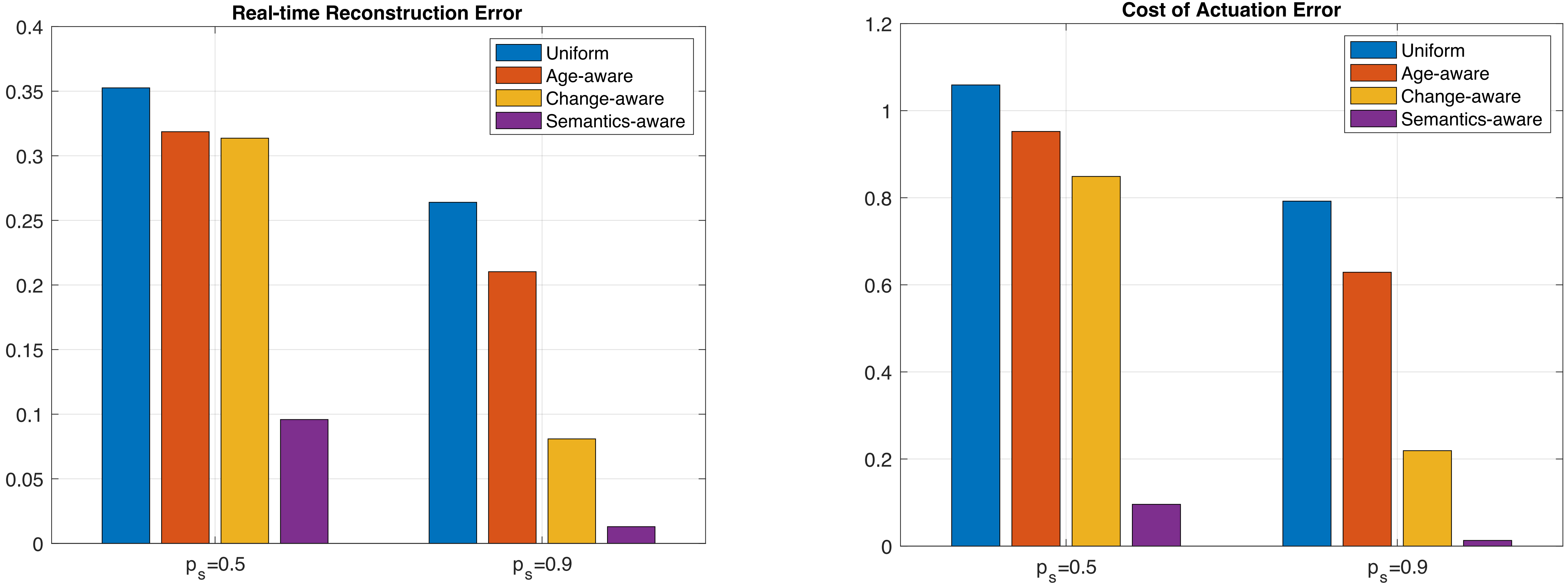
Sampling and communication policies

- Change-aware: sample generation is triggered at the transmitter whenever a change at the state of the source, with respect to the previous sample, is observed. (***No feedback or knowledge at the receiver's side required***)
- Semantics-aware: extends the Change-aware into that the amount of change is not solely measured at the source but is also tracked by the difference in state *between receiver and transmitter*.

- Sampling and transmission at every timeslot could provide the best performance for real-time reconstruction. It requires a very large number of samples, which are not necessarily useful and require excessive resources.
- The semantics-empowered policies *reduce or even eliminate* the generation of uninformative sample updates, thus *improving network resource usage*.

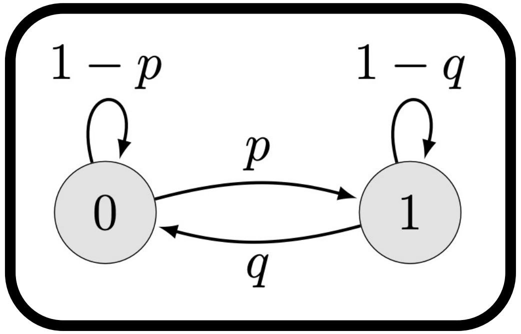
Slowly-varying source – ($p = 0.1$, $q = 0.15$)

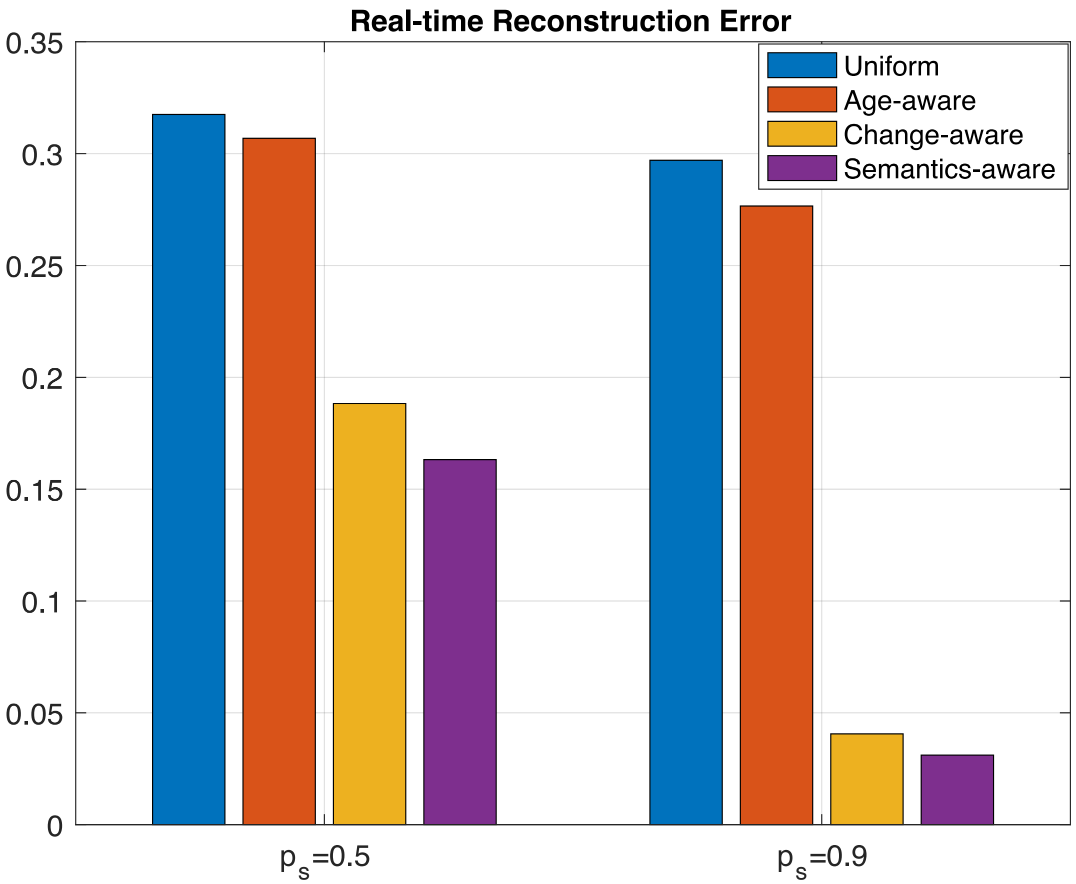


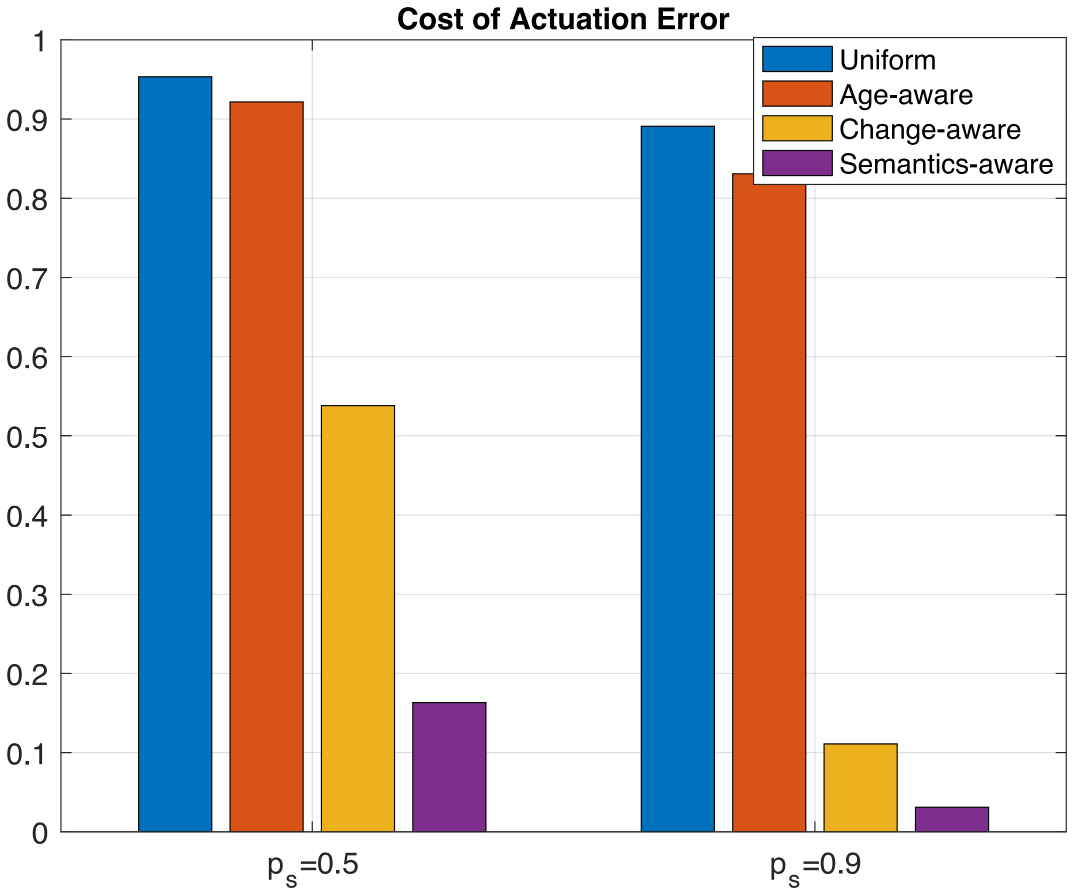


$C_{0,1}$=5, $C_{1,0}$=1

Rapidly-varying source – (p = 0.2, q = 0.7)







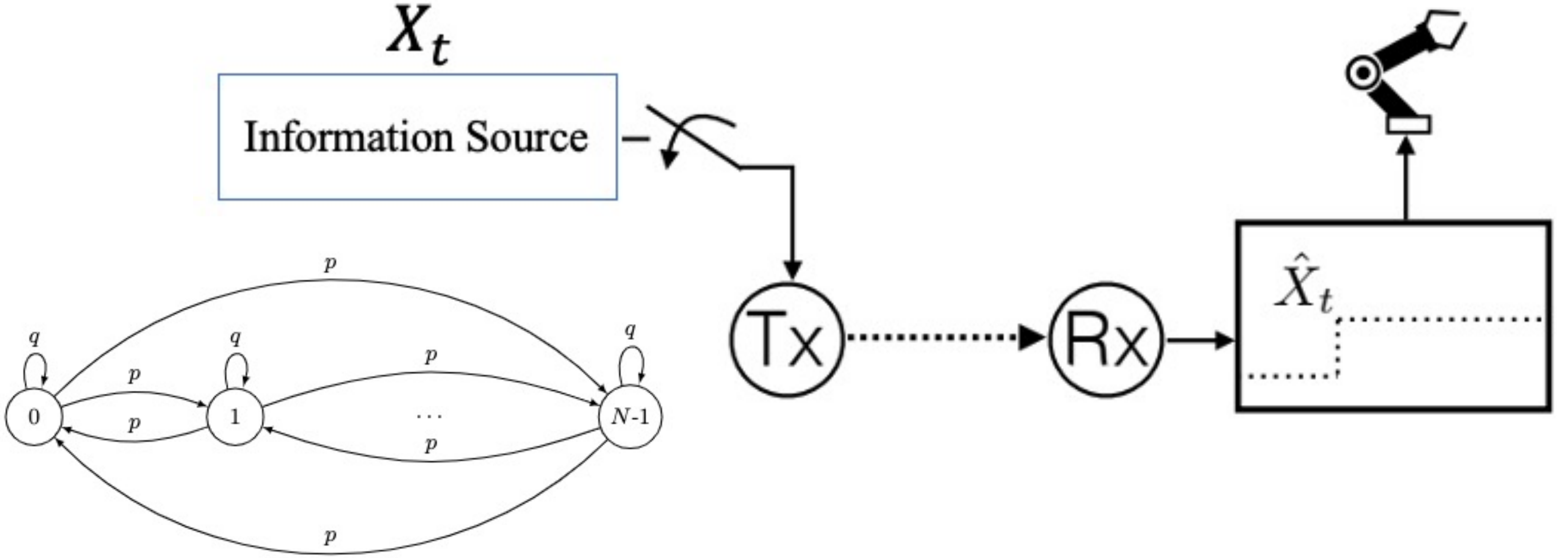
$C_{0,1}=5,\ C_{1,0}=1$

Real-time Remote Reconstruction of a Markov Source and Actuation over Wireless

S. Mehrdad, M. Kountouris, and N. Pappas, "*Real-time Remote Reconstruction of a Markov Source and Actuation over Wireless*", IEEE ICC Workshop on Semantic Communications, May 2023. arXiv:2302.01132, Feb. 2023 (longer version)

LINKÖPING
UNIVERSITY

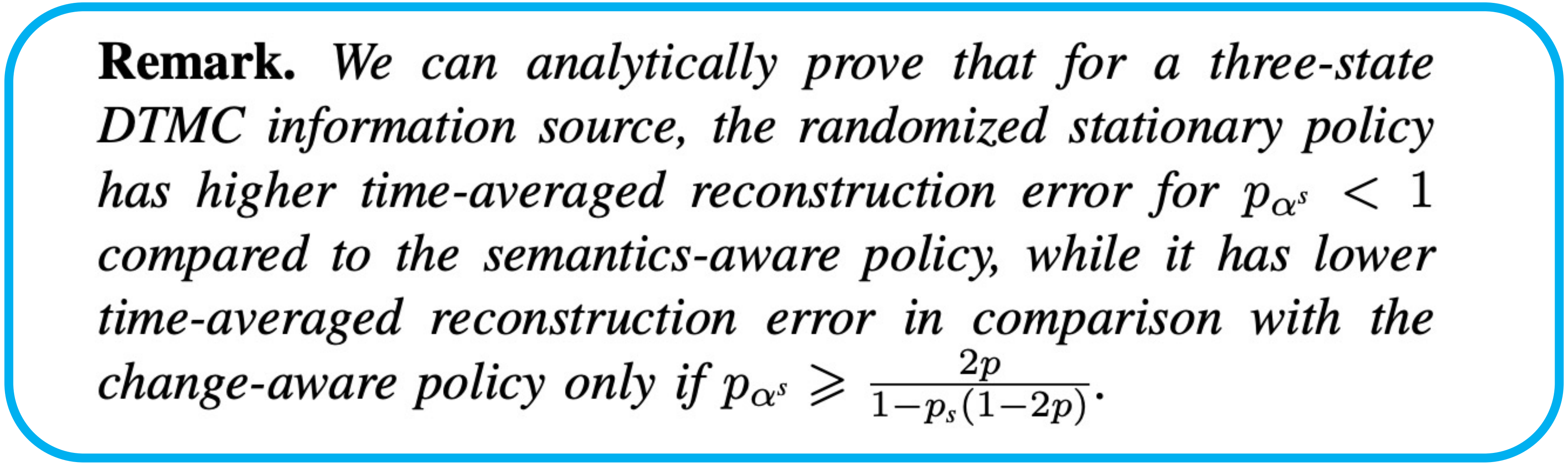
A more general model



N-state discrete time Markov chain (DTMC) to describe the information source/process evolution.

Randomized Stationary Policy

Sampling and transmission at each slot with *probability* p_{α^s}



Remark. *We can analytically prove that for a three-state DTMC information source, the randomized stationary policy has higher time-averaged reconstruction error for* $p_{\alpha^s} < 1$ *compared to the semantics-aware policy, while it has lower time-averaged reconstruction error in comparison with the change-aware policy only if* $p_{\alpha^s} \geqslant \frac{2p}{1-p_s(1-2p)}$.

Numerical results

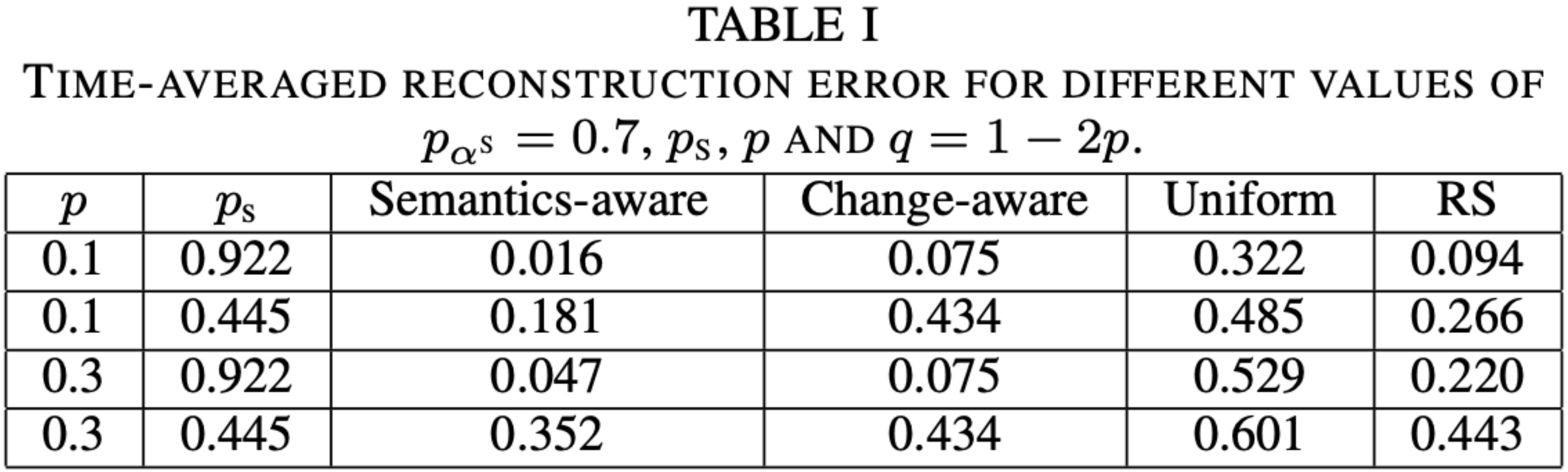


TABLE I

TIME-AVERAGED RECONSTRUCTION ERROR FOR DIFFERENT VALUES OF $p_{\alpha^{\mathrm{S}}} = 0.7$, p_{S}, p AND $q = 1 - 2p$.

p	p_{S}	Semantics-aware	Change-aware	Uniform	RS
0.1	0.922	0.016	0.075	0.322	0.094
0.1	0.445	0.181	0.434	0.485	0.266
0.3	0.922	0.047	0.075	0.529	0.220
0.3	0.445	0.352	0.434	0.601	0.443

Optimization

The cost of sampling at each attempted transmission

Total average sampling cost

$$\begin{aligned} &\underset{p_{\alpha^{\mathrm{s}}}}{\text{minimize}} && P_E \\ &\text{subject to} && \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\delta\mathbb{1}\{\alpha_t^{\mathrm{s}}=1\}\leqslant\delta_{\max}, \end{aligned}$$

Optimization

The cost of sampling at each attempted transmission

Total average sampling cost

$$\begin{aligned} \underset{p_{\alpha^{\mathrm{s}}}}{\text{minimize}} \quad & P_E \\ \text{subject to} \quad & \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \delta \mathbb{1}\{\alpha_t^{\mathrm{s}} = 1\} \leqslant \delta_{\max}, \end{aligned}$$

$$\Big\downarrow \text{N=2}$$

$$\begin{aligned} \underset{p_{\alpha^{\mathrm{s}}}}{\text{minimize}} \quad & \frac{2\big(p - p p_{\alpha^{\mathrm{s}}} p_{\mathrm{s}}\big)}{4p + 2p_{\alpha^{\mathrm{s}}} p_{\mathrm{s}} - 4p p_{\alpha^{\mathrm{s}}} p_{\mathrm{s}}} \\ \text{subject to} \quad & p_{\alpha^{\mathrm{s}}} \leqslant \eta. \end{aligned} \quad \longrightarrow \quad P_E^* = \frac{2\big(p - p p_{\mathrm{s}} \eta\big)}{4p + 2p_{\mathrm{s}}\eta - 4p p_{\mathrm{s}}\eta}.$$

$$\eta = \frac{\delta_{\max}}{\delta} \leqslant 1$$

Numerical results

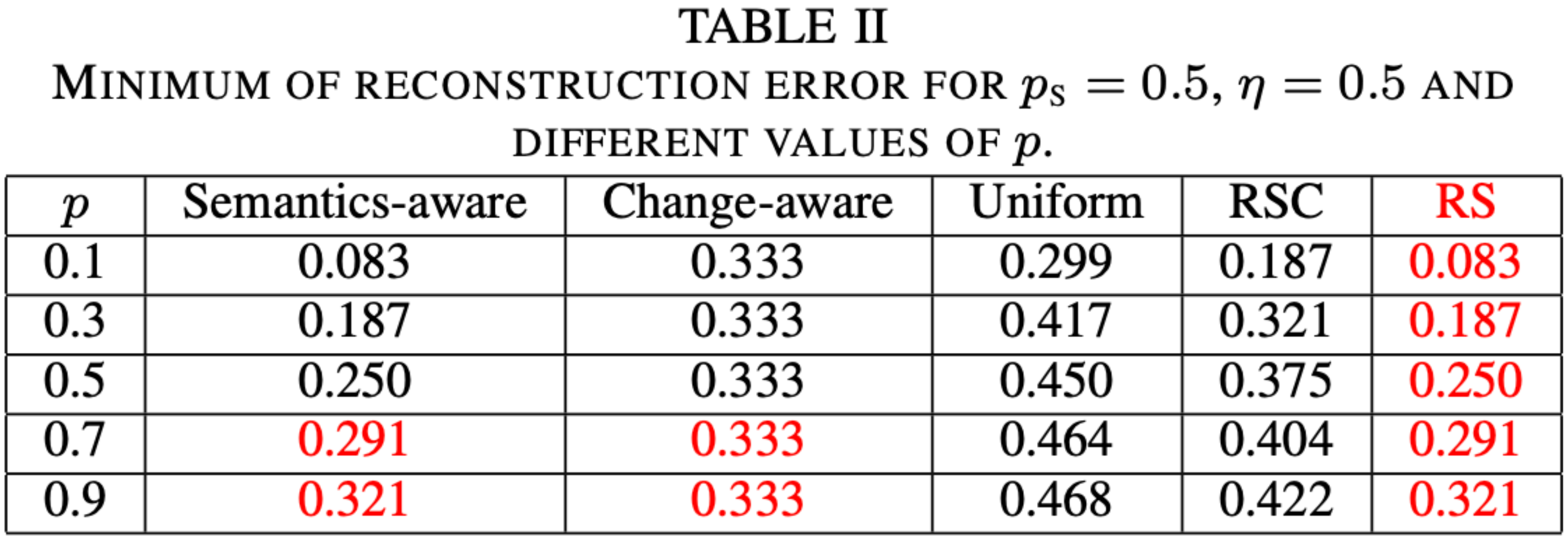


TABLE II

MINIMUM OF RECONSTRUCTION ERROR FOR $p_{\mathrm{S}} = 0.5$, $\eta = 0.5$ AND DIFFERENT VALUES OF p.

p	Semantics-aware	Change-aware	Uniform	RSC	RS
0.1	0.083	0.333	0.299	0.187	0.083
0.3	0.187	0.333	0.417	0.321	0.187
0.5	0.250	0.333	0.450	0.375	0.250
0.7	0.291	0.333	0.464	0.404	0.291
0.9	0.321	0.333	0.468	0.422	0.321

- The semantics-aware policy outperforms the optimal RSC policy when the source is slowly varying for

$$p \leqslant \frac{\eta p_{\mathrm{s}}}{1-2\eta+2\eta p_{\mathrm{s}}}$$

- In the unconstrained optimization, the optimal solution of RS is to generate w.p.1 samples, thus, it *generates an excessive amount of samples*.

Timing-aware error metrics

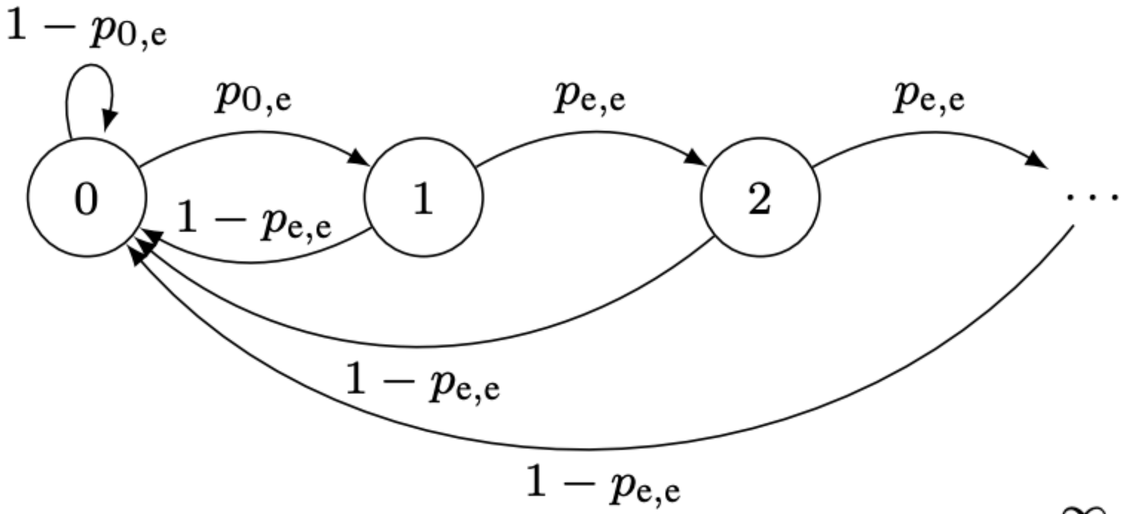
- In real-time and/or mission-critical applications being in an ***erroneous state*** for some time ***consecutively*** may lead to safety issues or could even have catastrophic consequences for the system.

- We propose
  - consecutive error metric
  - cost of memory error

Timing-aware error metrics – Consecutive Error

Consecutive Error (innate):

The number of consecutive time slots that system is in an erroneous state.



$$\pi_0 = \frac{1-p_{\mathrm{e,e}}}{1+p_{0,\mathrm{e}}-p_{\mathrm{e,e}}}$$

$$\pi_n = \frac{p_{0,\mathrm{e}}(1-p_{\mathrm{e,e}})p_{\mathrm{e,e}}^{n-1}}{1+p_{0,\mathrm{e}}-p_{\mathrm{e,e}}}$$

Average of Consecutive Error

$$\bar{C}_E = \sum_{x=1}^{\infty} x\pi_x = \frac{p_{0,\mathrm{e}}}{1+p_{0,\mathrm{e}}-2p_{\mathrm{e,e}}-p_{0,\mathrm{e}}p_{\mathrm{e,e}}+p_{\mathrm{e,e}}^2}$$

Timing-aware error metrics – Cost of Memory Error

- **Cost of memory error** (contextual)

$$C_{\mathrm{M}}(x) = \begin{cases} 0, & x = 0, \\ \kappa^{x}, & x = 1, 2, \cdots . \end{cases}$$

- *Penalization of memory error over n consecutive time slots*

$$\bar{C}_{E}^{\mathrm{M}} = \sum_{x=1}^{n} C_{\mathrm{M}}(x)\pi_{x} = \frac{\kappa p_{0,\mathrm{e}}(1 - p_{\mathrm{e,e}})\big(1 - (\kappa p_{\mathrm{e,e}})^{n}\big)}{(1 - \kappa p_{\mathrm{e,e}})\big(1 + p_{0,\mathrm{e}} - p_{\mathrm{e,e}}\big)}$$

Numerical results

- The semantics-aware policy exhibits smaller cost of memory error compared to all other policies.

- The semantics-aware policy does not allow the system to operate in an erroneous state for several consecutive time slots.

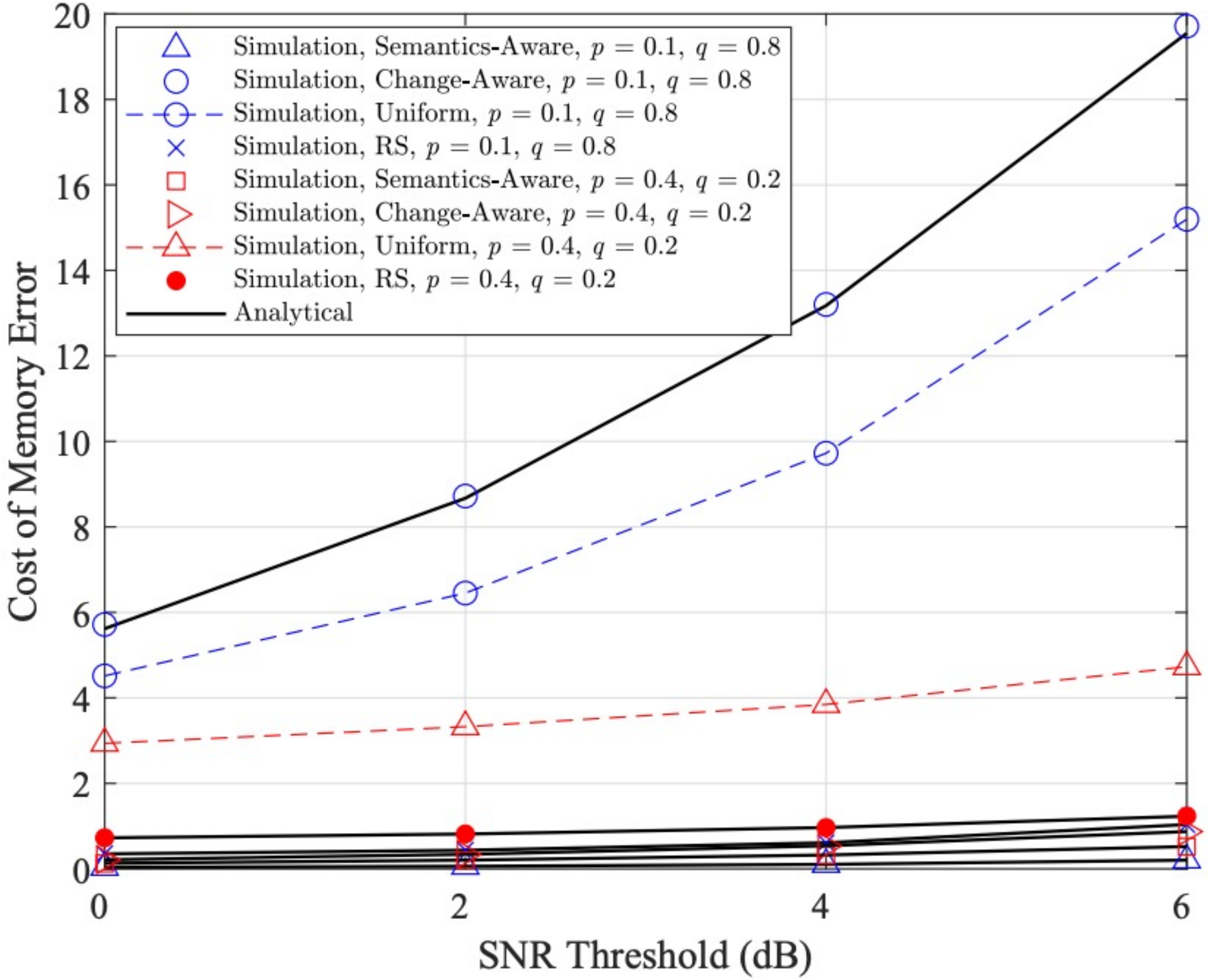


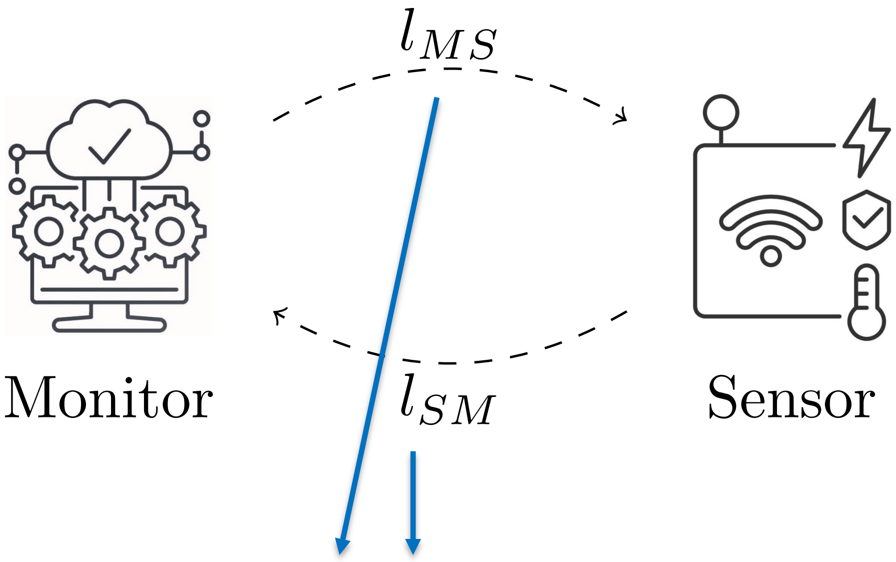
Fig. 3. Cost of memory error as a function of γ, for $n = 10$, $\kappa = 2$, $p_{\alpha^s} = 0.7$, and different values of p and q.

Fault Detection and autonomous maintenance

- **G. Stamatakis, N. Pappas, A. Fragkiadakis, A. Traganitis, "Semantics-Aware Active Fault Detection in IoT", 20th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), Sep. 2022.**
- G. Stamatakis, **N. Pappas**, A. Fragkiadakis, A. Traganitis, "Autonomous Maintenance in IoT Networks via AoI-driven Deep Reinforcement Learning", IEEE INFOCOM - 4th Age of Information Workshop, May 2021.

System Model

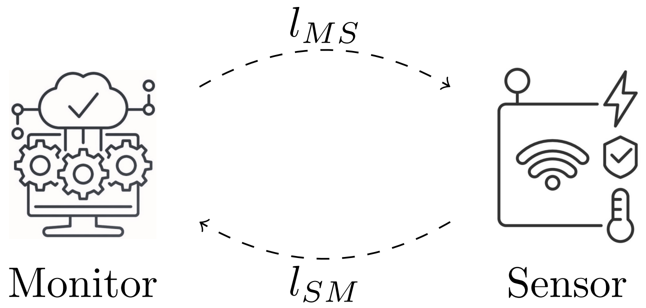
- A sensor transmits status updates to monitor.
- Wireless links are on/off channels according to a Markov Chain.
- If a link is in on/off state then the transmission is successful/failed.
- Slotted time.
- The sensor can be in faulty/healthy state according to a Markov Chain.
- If the sensor is in a healthy state can generate an update with probability P_g and transmits it over the link l_{SM}.



On/off channels according to a Markov Chain

After a successful reception of a probe through l_{MS}, the sensor will generate a fresh status update at the next time slot with probability 1/0, if it is in a healthy/faulty state.

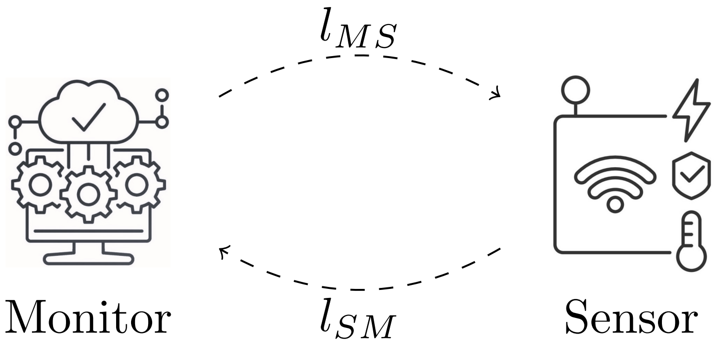
System Model



- *A monitoring agent must optimally decide, at the beginning of a time slot, to probe or not the sensor.*
- A transition cost is induced on the agent by the end of each time slot due to its decision and the dynamics of the system.
- The objective is to minimize the total accumulated cost over a finite time horizon.
- The transition cost is a function of
  - the agent's confidence in its belief about the joint health status of the sensor and links,
  - the *staleness of the status updates* it has received up to that time slot,
  - and a *cost value c* associated with the *probing action*.

Problem formulation

V_t quantifies the importance of receiving a fresh status update at the monitor at time t.



Value of Information

$$V_t = \lambda_1 H_t + \lambda_2 \bar{\Delta}_t$$

The agent’s confidence in the health status belief vector is expressed by its entropy.

Normalized AoI over the horizon

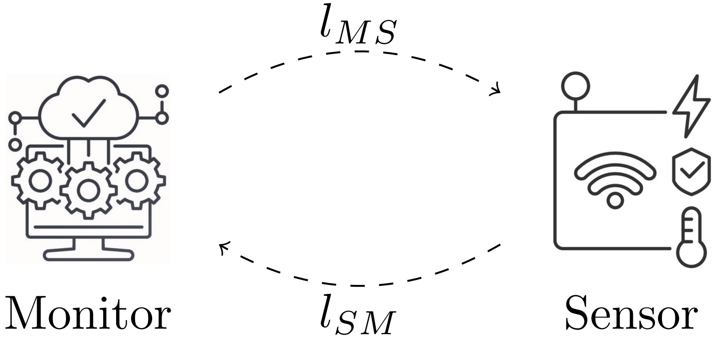
$$\Delta_t = \begin{cases} 1, & \text{if } z_t = 1 \\ \min\{N, \Delta_t + 1\}, & \text{if } z_t = 0 \end{cases}$$

Problem formulation

V_t quantifies the importance of receiving a fresh status update at the monitor at time t.

Value of Information

$$V_t = \lambda_1 H_t + \lambda_2 \bar{\Delta}_t$$



The agent’s confidence in the health status belief vector is expressed by its entropy.

Normalized AoI over the horizon

$$\Delta_t = \begin{cases} 1, & \text{if } z_t = 1 \\ \min\{N, \Delta_t + 1\}, & \text{if } z_t = 0 \end{cases}$$

Remark: Probing will lead to the reduction of both AoI and entropy?

Problem formulation

V_t quantifies the importance of receiving a fresh status update at the monitor at time t.

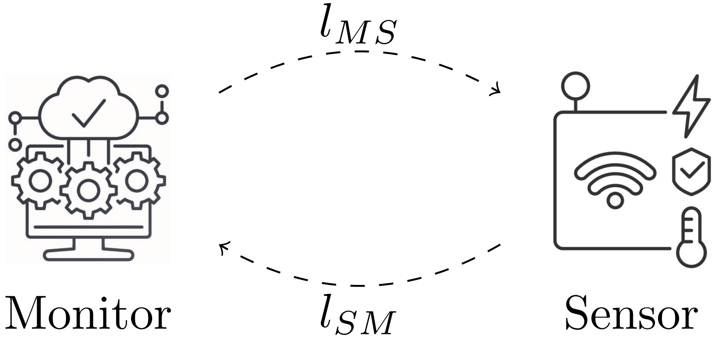
Value of Information

$$V_t = \lambda_1 H_t + \lambda_2 \bar{\Delta}_t$$

The agent's confidence in the health status belief vector is expressed by its entropy.

Normalized AoI over the horizon

$$\Delta_t = \begin{cases} 1, & \text{if } z_t = 1 \\ \min\{N, \Delta_t + 1\}, & \text{if } z_t = 0 \end{cases}$$



Remark: Probing will lead to the reduction of both AoI and entropy? **No, not always!**

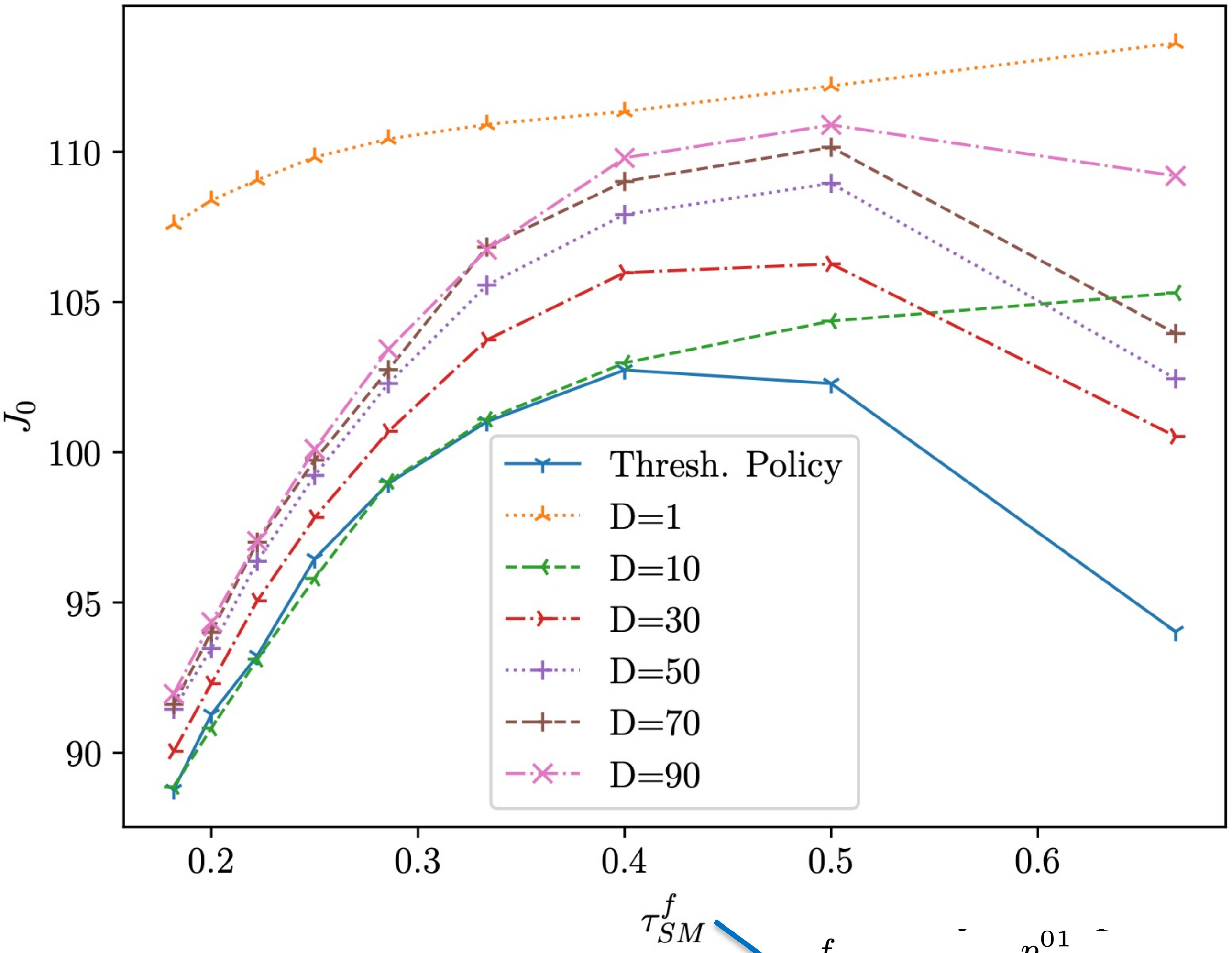
Probing makes the generation of a status update mandatory; however, probing introduces a new type of uncertainty due to the transmission failures.

Numerical results

$c = 1, \lambda_1 = 1, \lambda_2 = 1, P_g = 0.1, N=100$

$$P^{MS} = \begin{bmatrix} 0.9 & 0.1 \\ 0.9 & 0.1 \end{bmatrix} \quad P^{S} = \begin{bmatrix} 0.9 & 0.1 \\ 0.9 & 0.1 \end{bmatrix}$$

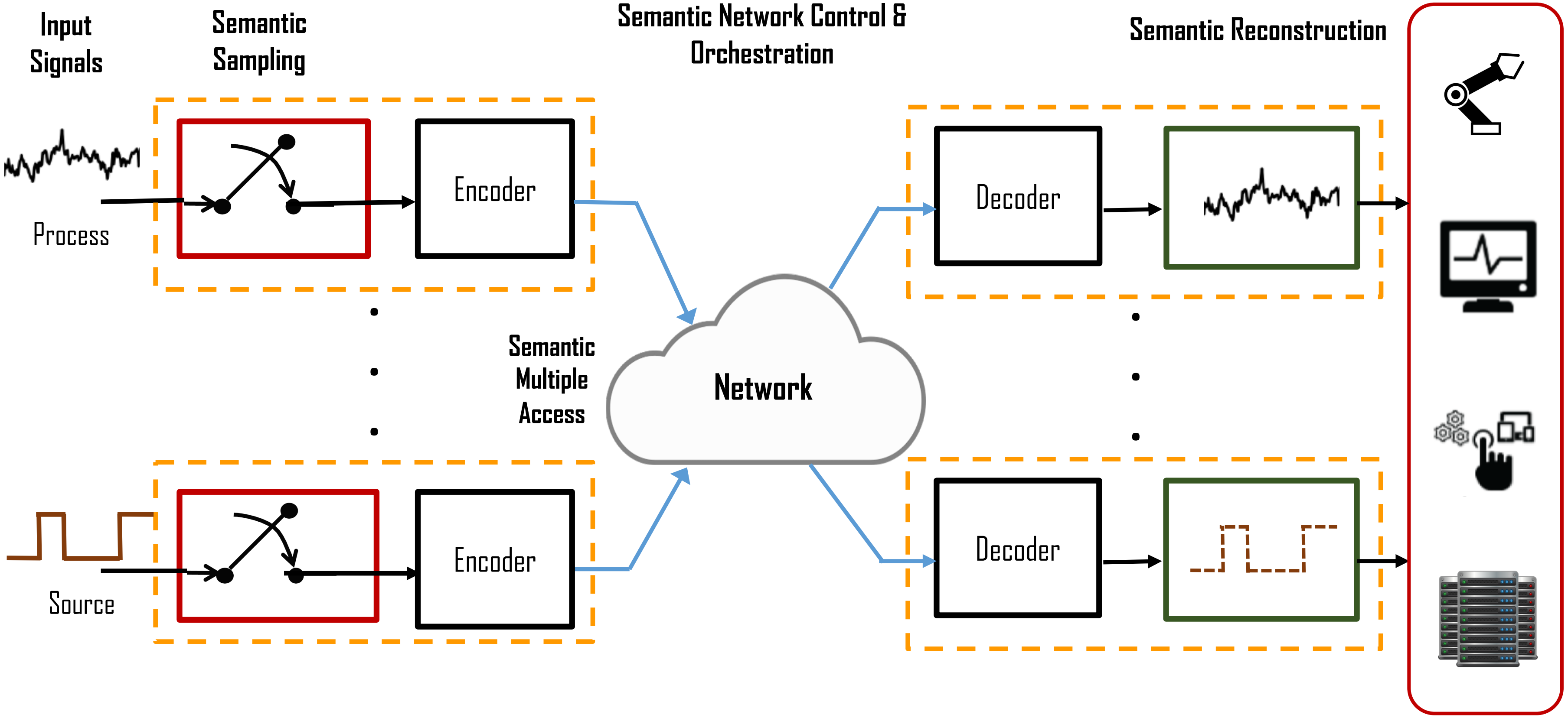
$$P^{SM} = \begin{bmatrix} 0.9 & 0.1 \\ 1 - p^{11}_{SM} & p^{11}_{SM} \end{bmatrix}$$



$$\tau^f_{SM} = \frac{p^{01}_{SM}}{1 - p^{11}_{SM} + p^{01}_{SM}}$$

- Comparative results with a delay based probing policy.
- For large values of τ^f_{SM} , the monitor was confident that l_{SM} was in a faulty state, and this resulted in a reduced cost due to health status entropy.
- Despite the reduction of J_o for all policies **the effect of persistent probing is still evident** and especially for the delay policies with D = 1 and 10.

Concluding remarks



- Communication process extends up to **goal-oriented signal reconstruction and information exploitation**
- A monitored signal: a physical phenomenon/event **distributed in space and evolving in time, space, and in its semantic (topological) space (as part of a larger one)!**

Projects funding this research

ETHOR

Co-funded by
the European Union

Swedish Research Council

ELLIIT

- "Semantics-Empowered Communication for Networked Intelligent Systems", Swedish Research Council
- "Information Handling in Industrial IoT", ELLIIT
- "Low Latency Communications for Wireless Networks: Exploiting Traffic Characteristics", CENIIT
- "Self-evolving terrestrial/non-terrestrial hybrid networks", HORIZON-JU-SNS-2022-STREAM-B-01-03 — Communication Infrastructure Technologies and Devices.
- *Two more Horizon Europe projects will start end of the year.*

Thank you!
Reach me at
nikolaos.pappas@liu.se

www.liu.se

LINKÖPING
UNIVERSITY