Quantifying the Merits of Network-Assist Online Learning in Optimizing Network Protocols

Xiangxiang Dai
The Chinese University of Hong Kong

Zhiyong Wang
The Chinese University of Hong Kong

Jiancheng Ye
The University of Hong Kong

John C.S. Lui
The Chinese University of Hong Kong

June 2024

Outline

- Introduction
- Online Learning Framework
- Algorithm Design
- Theoretical Analysis & Simulations
- Network Applications
- Conclusion

Leveraging MAB for Network Optimization

Adaptive Strategy:

- Optimizing network applications on decision-making.
- Online learning methods for unknown parameters.
- Multi-Armed Bandit (MAB) prominence in online learning.

MAB Framework:

- Decisions (arms): potential operating parameters.
- Payoff from unknown distributions.
- Continual adaptation through learned decisions.

Applications:

TCP Congestion Control (Nie et al. 2019).

Mobile Edge Caching (Yang et al. 2018).

Channel Selection (Liu et al. 2023).

Current Drawbacks and Our Goal

Static Assumptions of Basic MAB Models:

- Assume a static environment with payoff.
- Overlook the dynamics of real-world scenarios.

Limitations of Existing Approaches with State Detection:

- Via direct but inaccurate self-judgement or offline mappings.
- Network protocols may involve continuous arm spaces.

Our Objective:

Non-stationary online learning framework using network-assist signals for quick state change detection, enhanced learning in dynamic scenarios.

Contributions

Key Contributions:

1. **New Model:** Enhances network-assist online learning in non-stationary MAB environments.
2. **Algorithms:**
   - **DNS-UCB:** For discrete arm spaces.
   - **CNS-UCB:** For continuous arm spaces.
3. **Validation:** Theoretical analysis confirms efficiency, supported by simulations in both arm space types.
4. **Applications:** Cognitive radio network & rate-based TCP congestion control network applications.

System Model and Problem Formulation

Dynamic Environment and Network-Assist Signals:

- Time-varying states influenced by external factors like channel availability or network congestion
- *Network-assist* signals map states to corresponding arms.
- E.g., ECN reveals the current network congestion level, thus guiding to adjust the sending rate.

Roles of Network-Assist Signals:

- **State Reflection:** Reflect the current environmental condition.
- **Arm Guidance:** Direct to an arm subspace optimized for higher payoffs in current state.

Channel Selection and State Evolution

- State at each slot t denoted as s_t, transitions unknown in advance.
- Change-points g_i for each state change, with delay d_{g_i}.

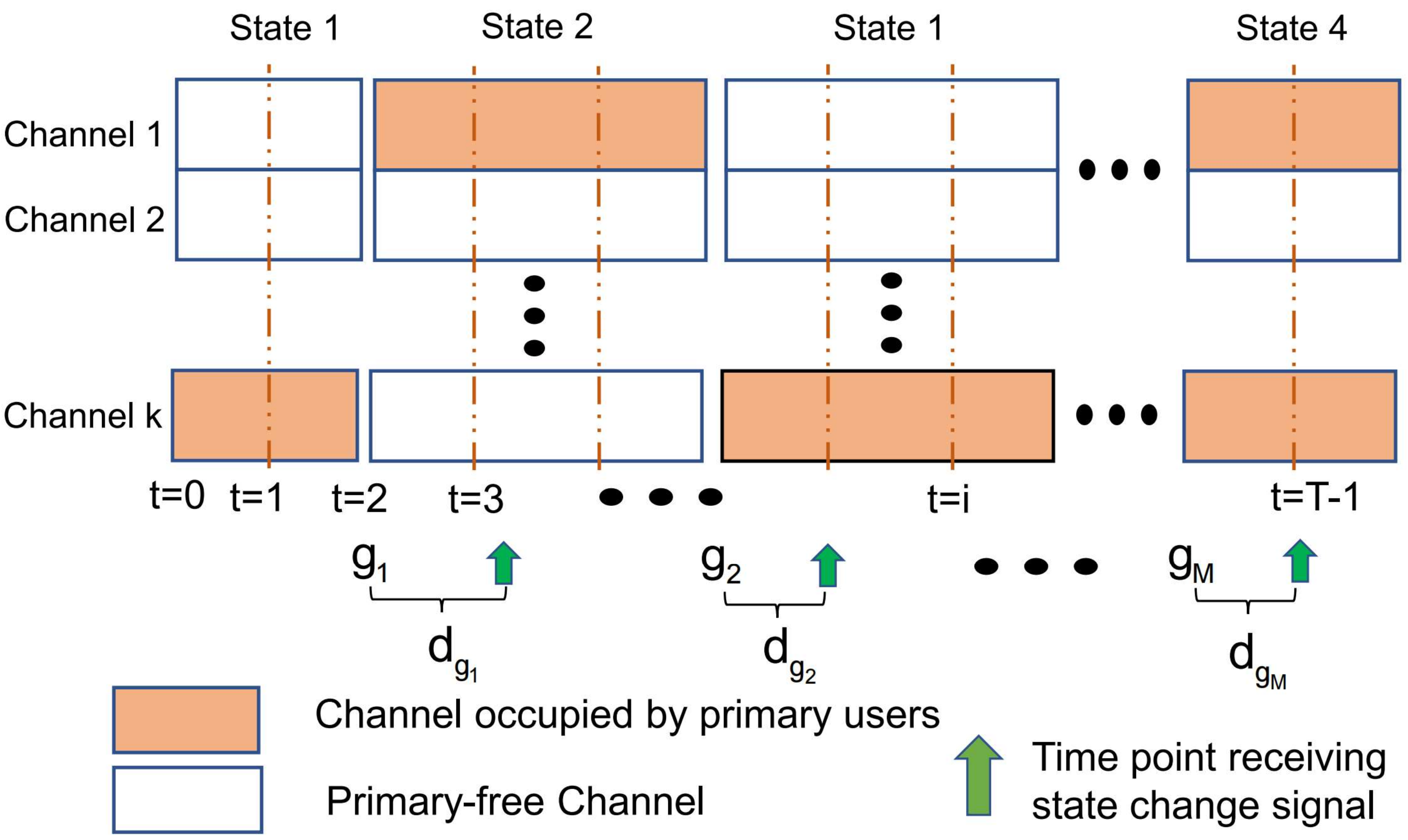

Figure 1: State evolution process in a wireless channel selection problem.

Payoff Distribution and Objective

State-Dependent Payoff Distributions:

- Each arm under state has a stationary distribution.
- Distribution variability by state affects utility, with normalized payoffs.

Learning Policy and Regret Minimization:

- Develop algorithms for discrete and continuous arm spaces to refine policy.

- Continuously adapt policy to minimize regret:

$$R_{\boldsymbol{p}}(T) = \mathbb{E}\left[\sum_{s\in\mathcal{S}} \sum_{t\in\mathcal{T}:s_t=s} \left(u_{s_t,t}(a^{\star}_{s_t}) - u_{s_t,t}(a_t^{p_{s_t}})\right)\right],$$

where $a^{\star}_{s_t}$ is the optimal arm for each state s_t.

Discrete Network-assist State-based UCB (DNS-UCB)

- Aimed at minimizing state-specific regrets.
- Decomposes the problem into manageable subroutines:
  - RCP (Remove-Check-Perform): Remove the wrong-state information stored dut to delay.
  - SUM (Select-Update-Maintain): Focus on long-term updates and maintenance of state policies.

Algorithm 1 DNS-UCB

1: Initialize $\mathcal{S}_0 = \emptyset$, $\mathcal{A}_s = \mathcal{A}$, $j = 1$; ▷ Initial setup
2: **for** all $t = 1, 2, \ldots, T$ **do**
3:     **if** Receive current network-assist state s_t **then**
4:         Update the arm subset $\mathcal{A}_{s_t}$; ▷ Update arms based on state
5:         Remove information, check if s_t appears and perform corresponding operations using Subroutine 1; ▷ Remove-Check-Perform
6:     **end if**
7:     Select an arm, update the statistical information, and maintain the sliding window using Subroutine 2; ▷ Select-Update-Maintain
8: **end for**

Discrete Network-assist State-based UCB (DNS-UCB)

State Management

- Arm subset from network signals upon state detection.

UCB-Based Arm Selection

- Exploitation of current information of empirically mean payoffs.
- Exploration for potential better arm not yet found.

Handling State Changes and Updates

- Continue with existing data if state revisited; restart if new.
- Update statistics with new payoff after arm selection.
- Eliminate incorrect data from mixed-state periods.

Continuous Network-assist State-based UCB (CNS-UCB)

```
Initialize S_0 = ∅, X_s = X, j = 1, H = 2;                                      ▷ Initial setup
while H ≤ T do                                                                   ▷ Double the horizon each iteration
    Set discretization parameter D;                                              ▷ Adjust discretization
    Determine the arm subset A_s of X_s;                                         ▷ Select appropriate arm subset
    for all t = H, H + 1, . . . , min(2H − 1, T) do                              ▷ Iterate over new horizon
        if Receive current network-assist state s_t then                         ▷ Check for new state
            Update the arm subspace X_{s_t};                                     ▷ Adjust arm subspace
            Obtain the subset A_{s_t};                                           ▷ Update arm subset
            Remove information, check if s_t appears and perform corresponding oper-
ations using Subroutine 1;                                                       ▷ Handle state change
        end if
        Select an arm, update the statistical information, and maintain the sliding
window using Subroutine 2;                                                       ▷ Operate selected arm
    end for
    H = 2H;                                                                      ▷ Expand time horizon
end while
```

Continuous Network-assist State-based UCB (CNS-UCB)

Introduction to CNS-UCB:

- Continuous space $\longrightarrow$ Discrete choice.

Technical Approach:

- Uniformly local Lipschitz condition to guide discretization with discretization parameter D.

Algorithm Structure:

Outer Loop: Scales H iteratively using a doubling strategy.

Inner Loop: Adapts to state changes by selecting subspaces.

Performance Analyis

Theorem 1

The regret upper bound for DNS-UCB is:
$O\left(\sqrt{|\mathcal{S}||\mathcal{A}|_{max} T \log(T)}\right) + O\left(CM\right)$, *where* $|\mathcal{A}|_{max} = \max_{s\in\mathcal{S}} |\mathcal{A}_s|$.

Theorem 2

The regret upper bound for CNS-UCB is:
$O\left(\left(L|\mathcal{S}|^{\frac{\beta}{2\beta+1}} + |\mathcal{S}|^{\frac{1}{2}}\right) T^{\frac{\beta+1}{2\beta+1}} \log^{\frac{\beta}{2\beta+1}}(T)\right) + O\left(CM\right)$, *with* $\mathcal{X} = [0,1]$, *and memory requirement* $\tilde{O}\left(T^{1/2\beta+1}\right)$.

Simulation Setup and Comparative Analysis

DNS-UCB Simulation Setup:

1. States and Arms: 4 states, each with subset of 12 arms.
2. Payoff Model: Payoff follows a Bernoulli distribution.
3. Maximum delay $C = 500$ time slots.

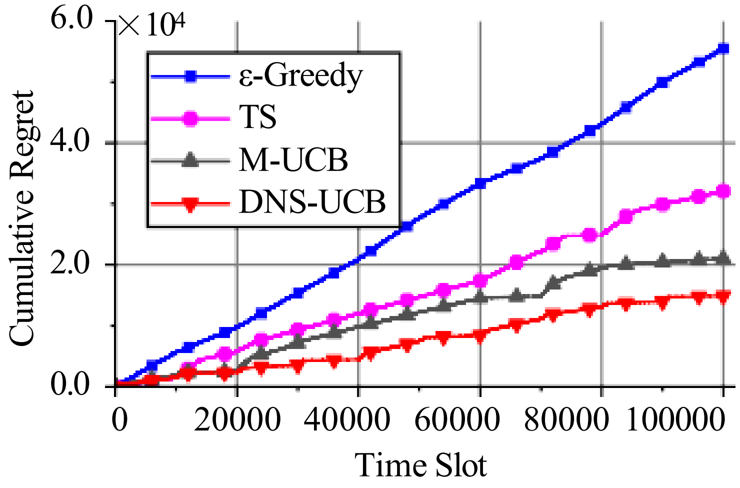

Figure 2: Discrete Arm Space

CNS-UCB Simulation Setup:

1. Continuous arm space range set to $[1, R]$, with $R = 8000$.
2. Dynamic payoff functions:
   - Even states: $\bar{u}_s(a) = -(a - \frac{1}{8}R)^2$
   - Odd states: $\bar{u}_s(a) = -(a - \frac{3}{4}R)^2$

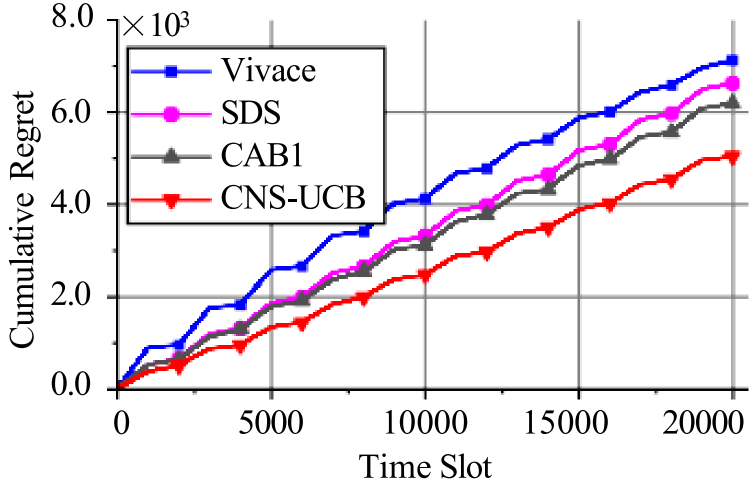

Figure 3: Continuous Arm Space

Wireless Network Channel Selection

User Types and Access:

- Primary Users (PUs): Licensed priority access.
- Secondary Users (SUs): Opportunistic, non-interfering access.

Challenge and Goal:

- Precise idle spectrum detection needed for SU coexistence with PUs.
- Enable efficient spectrum utilization by SUs.

Operational Dynamics:

- Spectrum Mobility: Changes with PUs' presence, affecting state.
- Spectrum Sensing: Identify state s_t with "network-assist" after delays.
- Spectrum Decision: SUs use learned policies to select channels.

Simulation Setup and Results Analysis

Settings Overview:

- Data: 4G channel throughput trace.
- Channels: 12 or 24.
- States: 4 or 8.
- Throughput reduction factor for PU-occupied channels.

Results Summary:

- DNS-UCB shows lower regret vs. M-UCB and TS.
- Stable and high throughput in DNS-UCB series.
- Efficient time usage by DNS-UCB in dynamic scenarios.

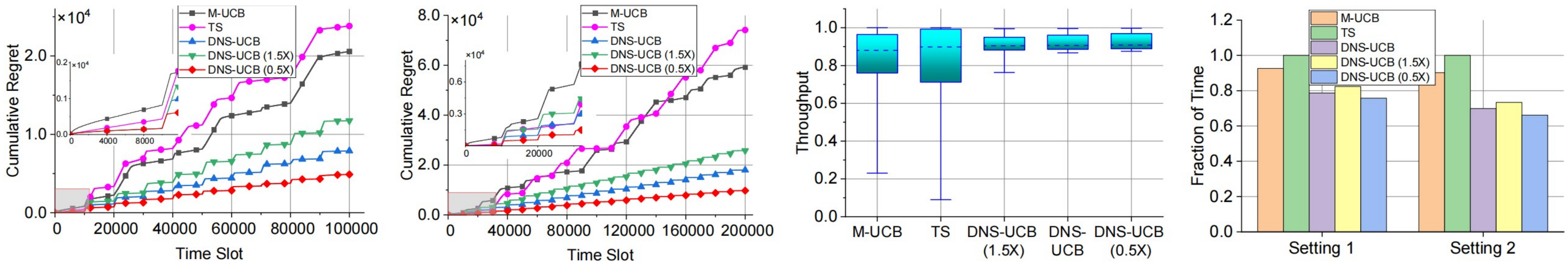

Figure 4: Comparisons under the wireless network channel selection problem.

Rate-Based TCP Congestion Control

Context and Motivation:

- BBR Protocol: We focus on intial phases—Startup.
- Emergence of Short Flows: Increasing prevalence, often completed within a few RTTs, challenges in the Startup phase.

Adjusting BBR for Dynamic Scenarios:

- HighGain Parameter: Critical in Startup; traditionally fixed but suboptimal for varying network conditions.
- CNS-UCB Approach: Dynamically selects optimal HighGain.

Problem Mapping

- ECN as Network-Assist: Utilizes ECN ratio from RED-AQM scheme.
- Reward Function: $r_a(t) = 1 - e_a(t)$, lower ECN ratio, less congestion.
- Arm Space: HighGain range in [2, 3] from the initial fixed value.

Results Analysis

Overview of Network Performance:

- BBR shows reduced throughput.
- **CNS-BBR** enhances pacing rates.
- **Long Flows:** Improvement in pacing rates during peak congestion times (12s to 18s).
- **Short Flows:** without compromising the network efficiency.

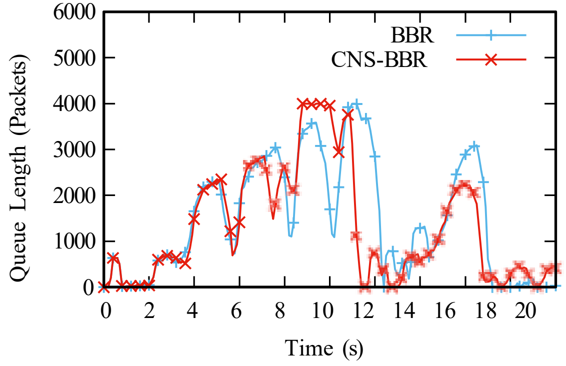

Figure 5: Queue Length at Router A.

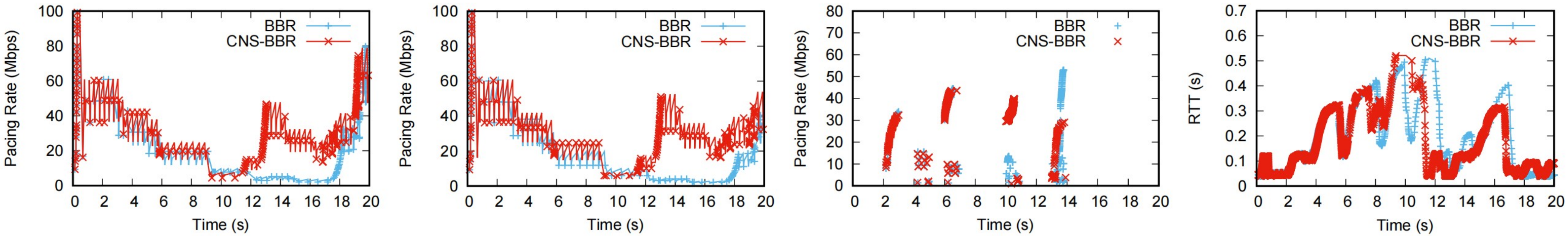

Figure 6: Comparisons for the rate-based TCP congestion control scenario.

Conclusion

Introduction:

- Pioneering network-assist online learning for dynamic environments.

Key Ponit:

- Network-assist signals optimized for dynamic conditions across discrete and continuous spaces.

Algorithm Development:

- DNS-UCB and CNS-UCB, with *sub-linear* optimal regret bounds.

Applications and Impact:

- Improve channel selection in cognitive radio network & rate-based congestion control in TCP.

Validation:

- Numerical & packet-level ns-3 simulations .

☺Thank You for Listening!

Feel free to reach me at
xiangxdai0@gmail.com

References

Xutong Liu, Jinhang Zuo, Hong Xie, Carlee Joe-Wong, and John CS Lui. Variance-adaptive algorithm for probabilistic maximum coverage bandits with general feedback. In *IEEE INFOCOM 2023-IEEE Conference on Computer Communications*, pages 1–10. IEEE, 2023.

Xiaohui Nie et al. Dynamic tcp initial windows and congestion control schemes through reinforcement learning. *IEEE Journal on Selected Areas in Communications*, 37(6):1231–1247, 2019.

Peng Yang, Ning Zhang, Shan Zhang, Li Yu, Junshan Zhang, and Xuemin Shen. Content popularity prediction towards location-aware mobile edge caching. *IEEE Transactions on Multimedia*, 21(4):915–929, 2018.