Multi-agent Traffic Load Balancing by Agents with Two-Layer Control Plane

E.P. Stepanov, R.L. Smeliansky

Lomonosov Moscow State University

11th International Conference "Distributed Computing and Grid Technologies in Science and Education" (GRID'2025)

July 8, 2025

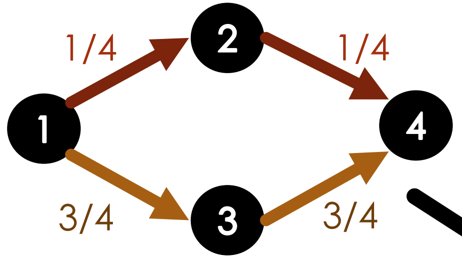
LOAD BALANCING PROBLEM STATEMENT

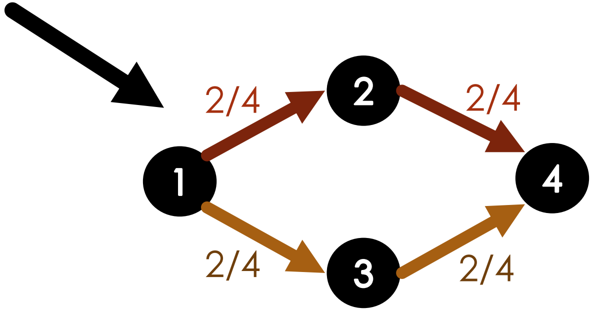
Objective function: $\Phi = \frac{1}{|E|} \sum_{(u,v)\in E} \left(\frac{b_{u,v}(X)}{c_{u,v}} - \mu'(X)\right)^2$

Average channel load: $\mu'(X) = \frac{1}{|E|} \sum_{(u,v)\in E} \frac{b_{u,v}(X)}{c_{u,v}}$

Example:

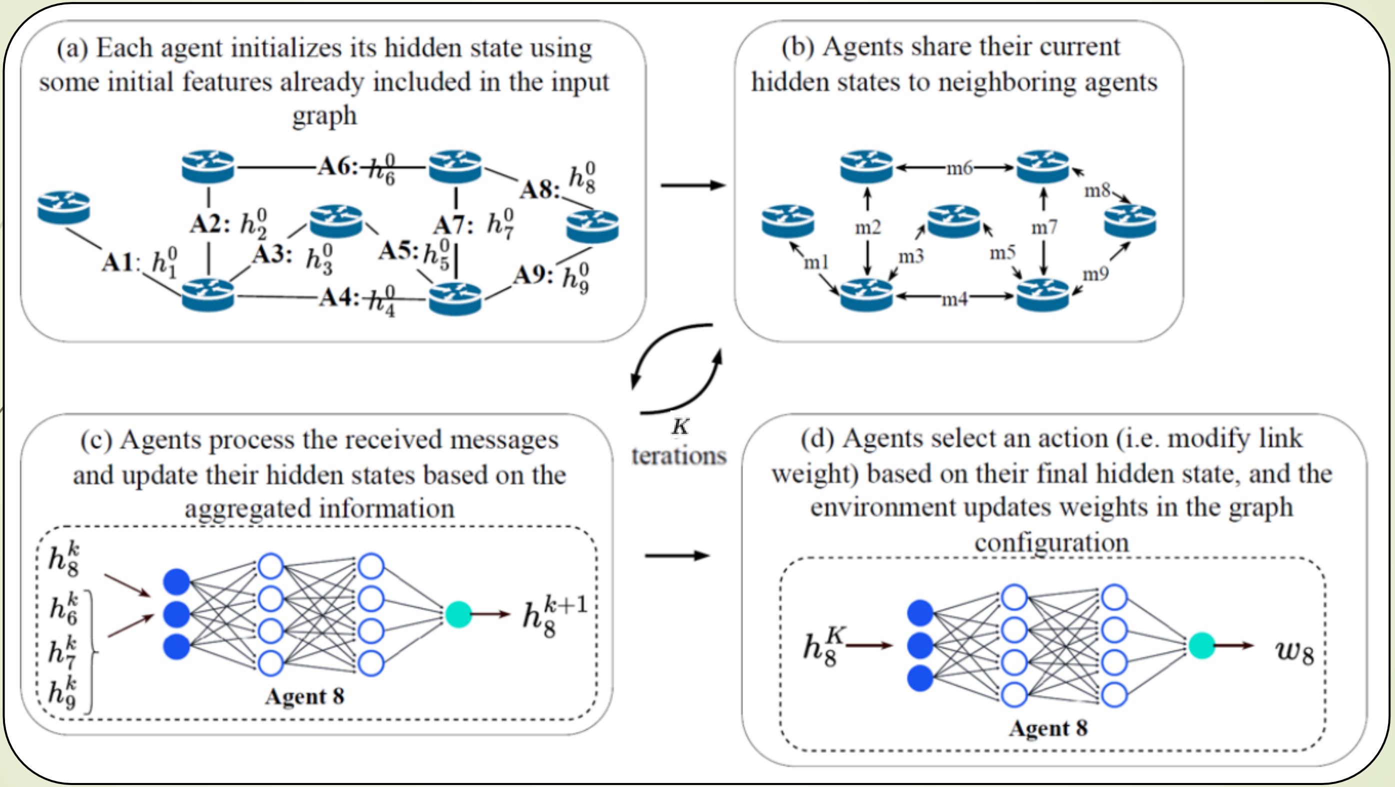
$\{f_1 = (1,4,1),\ f_2 = (1,4,1),\ f_3 = (1,4,1),\ f_4 = (1,4,1)\}$





- $G = (N, E)$ - directed graph:
  - N – set of switches.
  - E – set of (overlay) channels.
- For $(\mathrm{u}, \mathrm{v}) \in E$:
  - $b_{\mathrm{u,v}}$ - maximal bandwidth.
  - $c_{\mathrm{u,v}}$ - occupied bandwidth.
  - $c_{\mathrm{u,v}} / b_{\mathrm{u,v}}$ - channel load.
- Each flow is determined by its source, destination and rate – (f_s, f_d, r_f).
- Find weights $w_{\mathrm{u,v}}$ to minimize Φ.

MAROH – Multi-Agent ROuting with Hashing



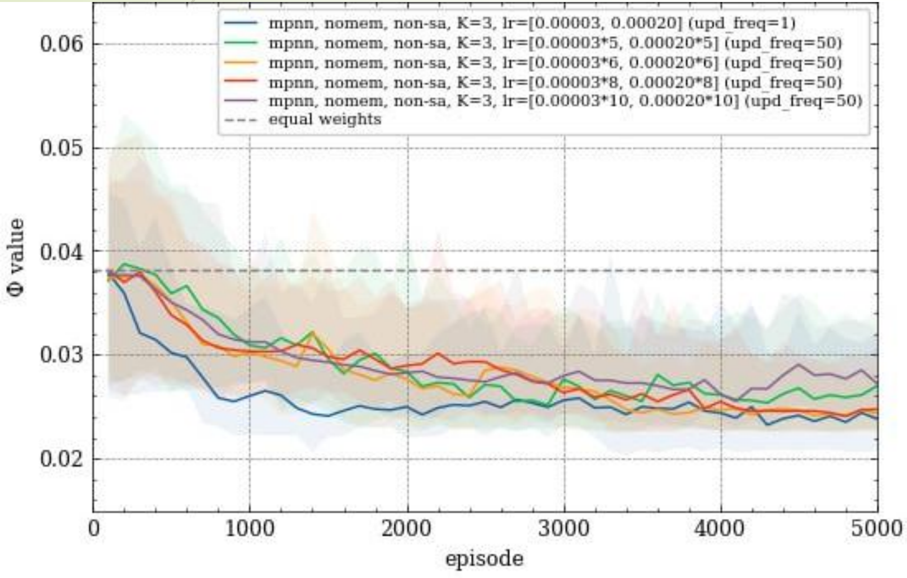
MAROH: Each weight update requires K **iterations of hidden states exchange**

SAMAROH - SIMULTANEOUS ACTIONS MAROH

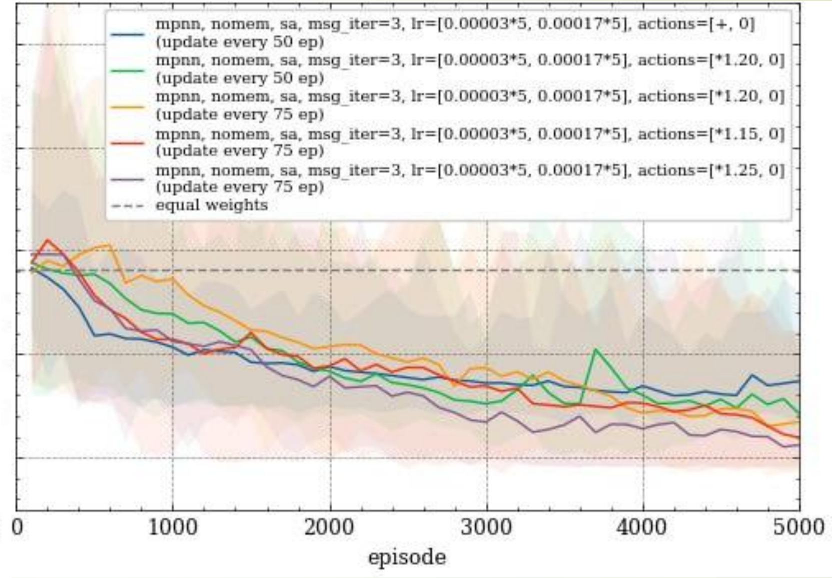
The original MAROH assumed that only a single agent could act simultaneously to ensure theoretical convergence guarantees.

What if we relax this constraint? Will the system still converge with parallel agent executions?

MAROH

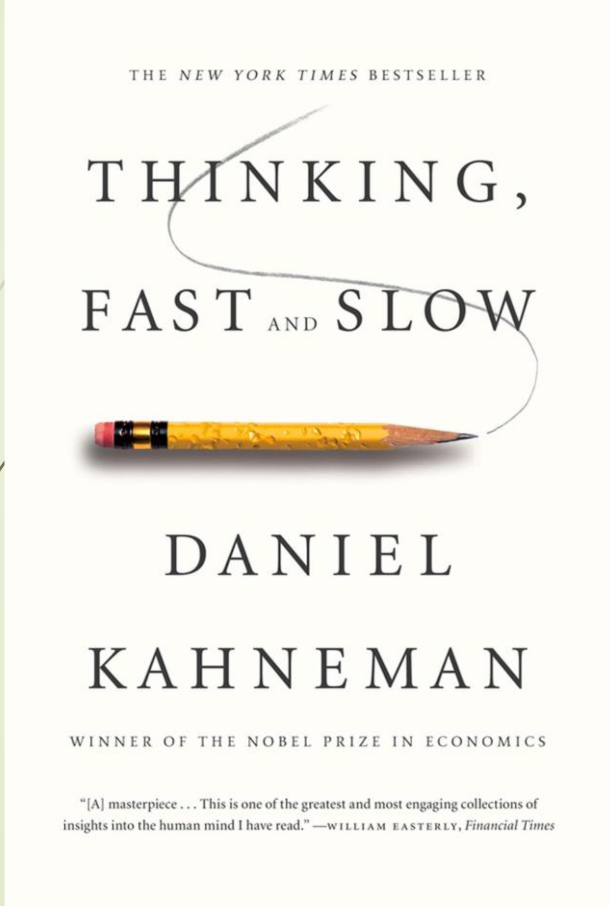


SAMAROH



SAMAROH preserves solution quality while cutting communication costs by 80%.

THINKING, FAST AND SLOW



Kahneman describes two different ways the brain forms thoughts:

System 1: Fast, automatic, frequent, emotional, stereotypic, unconscious. Examples of things system 1 can do:

- determine that an object is at a greater distance than another
- localize the source of a specific sound
- complete a common phrase (e.g. "war and ...")
- solve basic arithmetic (e.g. 2 + 2 = ?)

System 2: Slow, effortful, infrequent, logical, calculating, conscious. Examples of things system 2 can do:

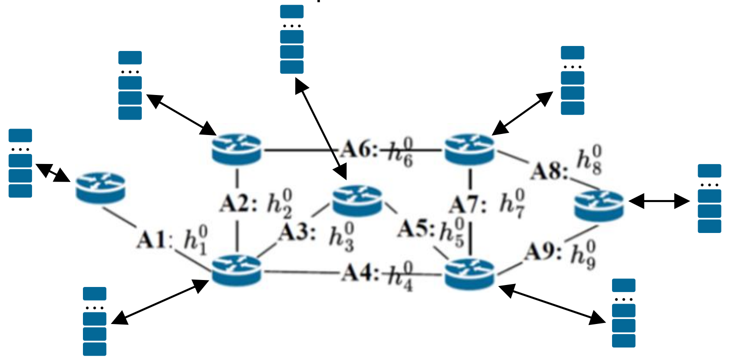
- look for a person with a particular feature
- try to recognize a sound
- park into a tight parking space
- determine the validity of a complex logical reasoning
- multiply two-digit numbers (e.g. 17 × 24)

TWO-LAYER CONTROL PLANE MAROH

Experience layer (Fast)

In familiar scenarios, where agents can rely solely on their own learned experience, communication becomes redundant, enabling immediate action execution

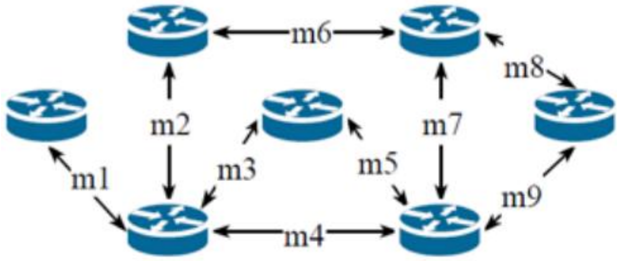
Agents act based on their collected experience



Communication layer (Slow)

In unfamiliar scenarios, MAROH agents exchange hidden states through communication to collaboratively determine optimal actions

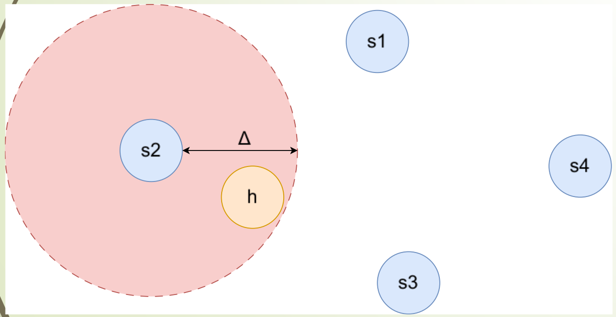
Agents share their current hidden states to neighboring agents

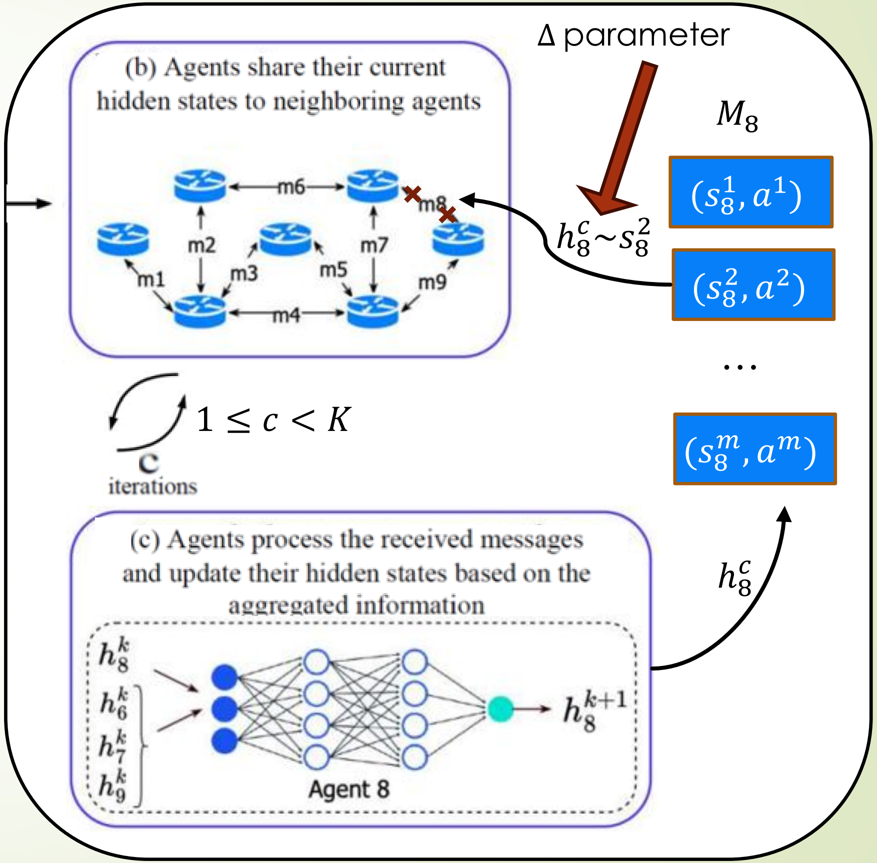


- How can we quantitatively distinguish "familiar" vs. "unfamiliar" states?
- What algorithmic strategies enable resource-constrained agents to selectively forget less relevant experiences?

EXPERIENCE LAYER

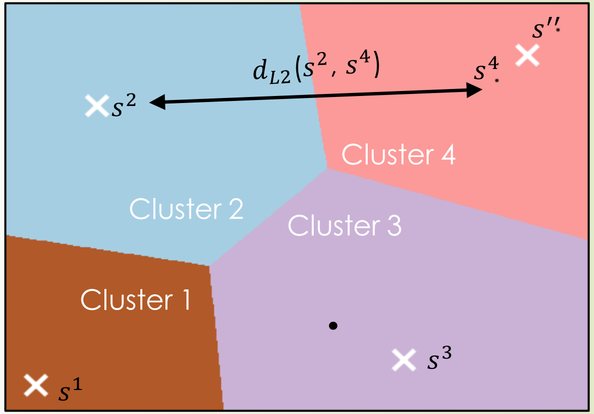
- Agent's memory is an array, consisting of m pairs (states and actions):
  - $M_i = [(s_i^j, a^j), j = 1 \dots m], i = 1 \dots n$
- New agent hidden states are matched to the closest memory state by the following distance metrics:
  - Euclidean: $d_{L2}(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$
  - Manhattan: $d_{L1}(x, y) = \sum_{i=1}^{n}|x_i - y_i|$
  - Cosine: $d_{cos}(x, y) = \frac{x*y}{\|x\|\|y\|} = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}}$





SELECTIVE STATE RETENTION VIA CLUSTERING-BASED FORGETTING

- The clustering algorithm identifies and stores only the most distant (centroid) states, with all other states within a cluster being represented by their nearest centroid.
- The clustering algorithm should satisfy the following requirements:
  - Hyperparameters should include the number of clusters;
  - Euclidean and other metrics support;
  - Implementation of algorithm supports more than 4096 clusters;



Clustering algorithm	K-Means	Agglomerative Clustering	Mini-Batch K-Means
Two-layer algorithm execution time (sec)	1.43	0.17	0.06
Communication reduction (%)	91.09	90.08	60.19

COMPARATIVE ANALYSIS

<table>
<tr><th>Algorithm</th><th>Clustering method</th><th>Metric</th><th>Number of Horizons</th><th>Φ Mean value</th><th>Number of exchanges (%)</th></tr>
<tr><td>MAROH</td><td colspan="2" rowspan="2"></td><td></td><td>0.203</td><td>100</td></tr>
<tr><td>SAMAROH</td><td rowspan="3">20</td><td>0.157</td><td>20</td></tr>
<tr><td rowspan="2">Two-layer SAMAROH</td><td>Agglomerative Clustering</td><td>COS</td><td>0.151</td><td>17.27</td></tr>
<tr><td rowspan="2">MiniBatch KMeans</td><td rowspan="2">L2</td><td>0.152</td><td>12.68</td></tr>
<tr><td>Two-layer MAROH</td><td>100</td><td>0.192</td><td>53.73</td></tr>
</table>

Two-layer control plane Simultaneous Actions MAROH (SAMAROH) showed **better objective function value** while **reducing number of exchanges more than 87%.**

CONCLUSION

- **Challenge**: reducing message exchanges is crucial for faster decisions and efficient resource use in MARL.
- **Solution 1**: SAMAROH enables simultaneous actions, cutting iterations but increasing convergence time.
- **Solution 2**: MAROH's two-layer approach memorizes familiar scenarios, requiring careful tuning of familiar state recognition thresholds, memory capacity, distance metrics
- **Result**: Combined methods reduce exchanges by **87%** while preserving objective function value.

Thank you for your attention!