The Edward S. Rogers Sr. Department
of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

Multi-Agent Deep Reinforcement Learning for Cooperative Edge Caching via Hybrid Communication

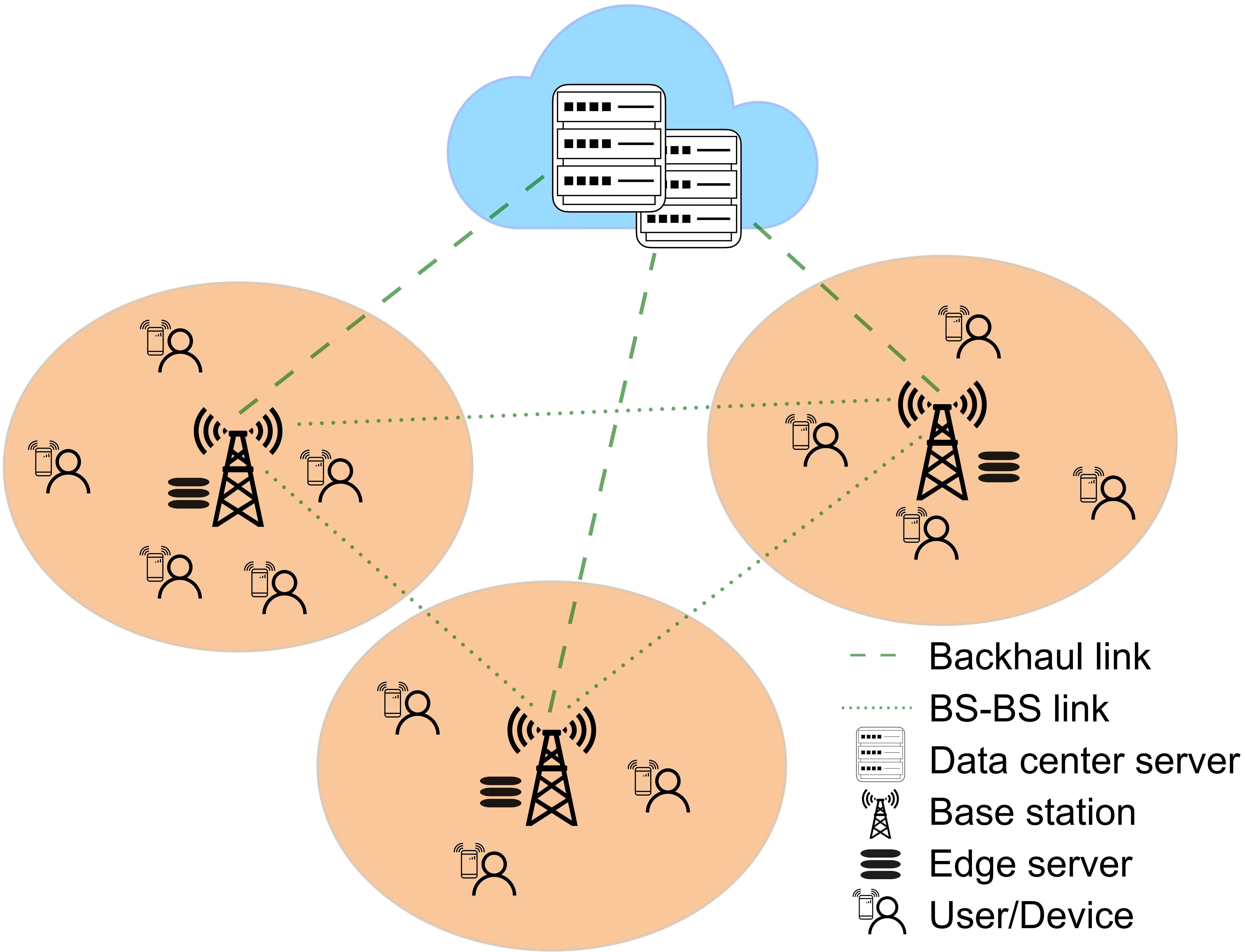
Fei Wang[1], Salma Emara[1], Isidor Kaplan[1], Baochun Li[1], Timothy Zeyl[2]
[1]Department of Electrical and Computer Engineering, University of Toronto
[2]Huawei Canada

ICC'23

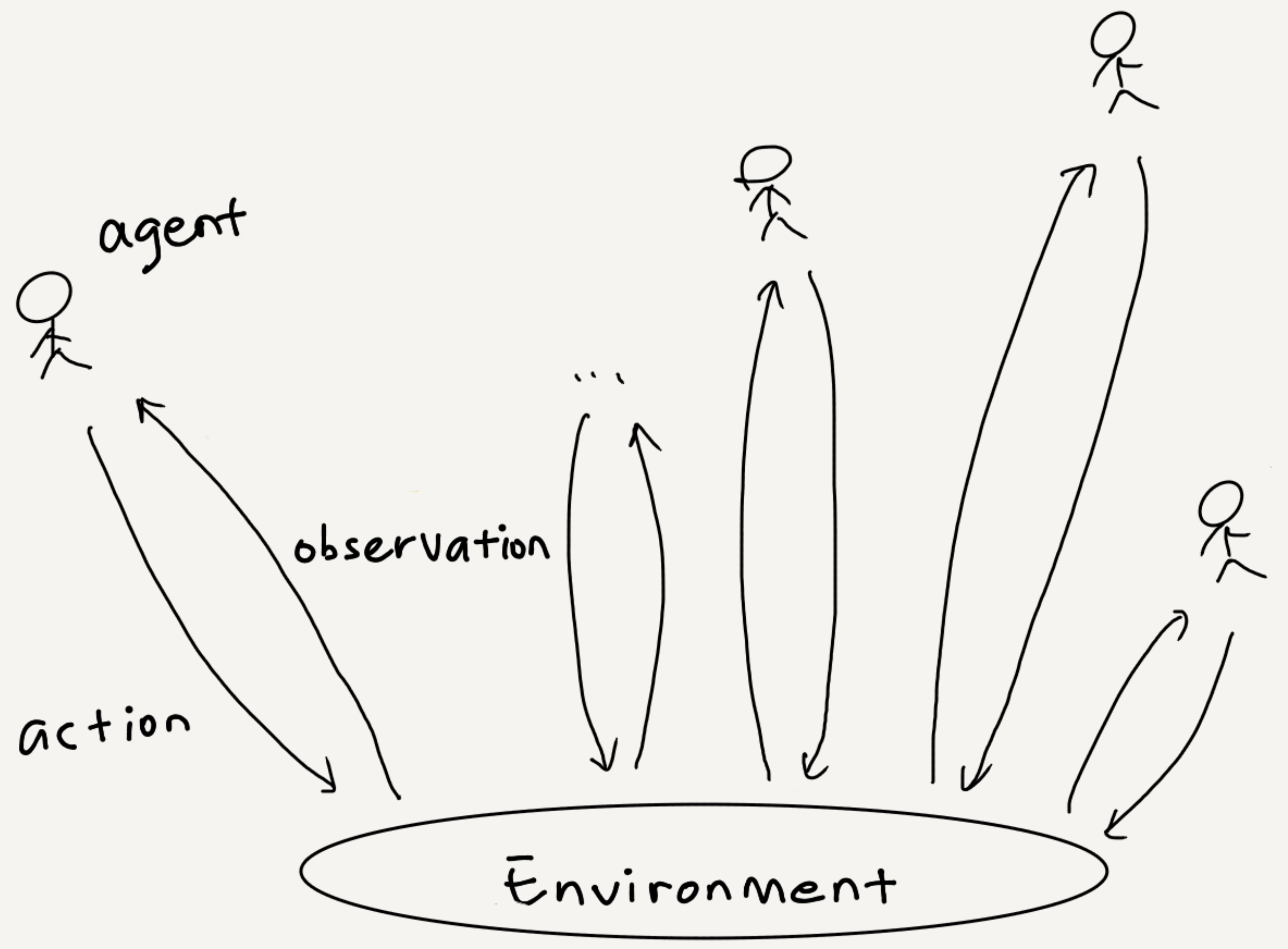
Outline

- Background
- Related Work
- Objective
- Design
- Experimental Results
- Conclusion

Cooperative Edge Caching

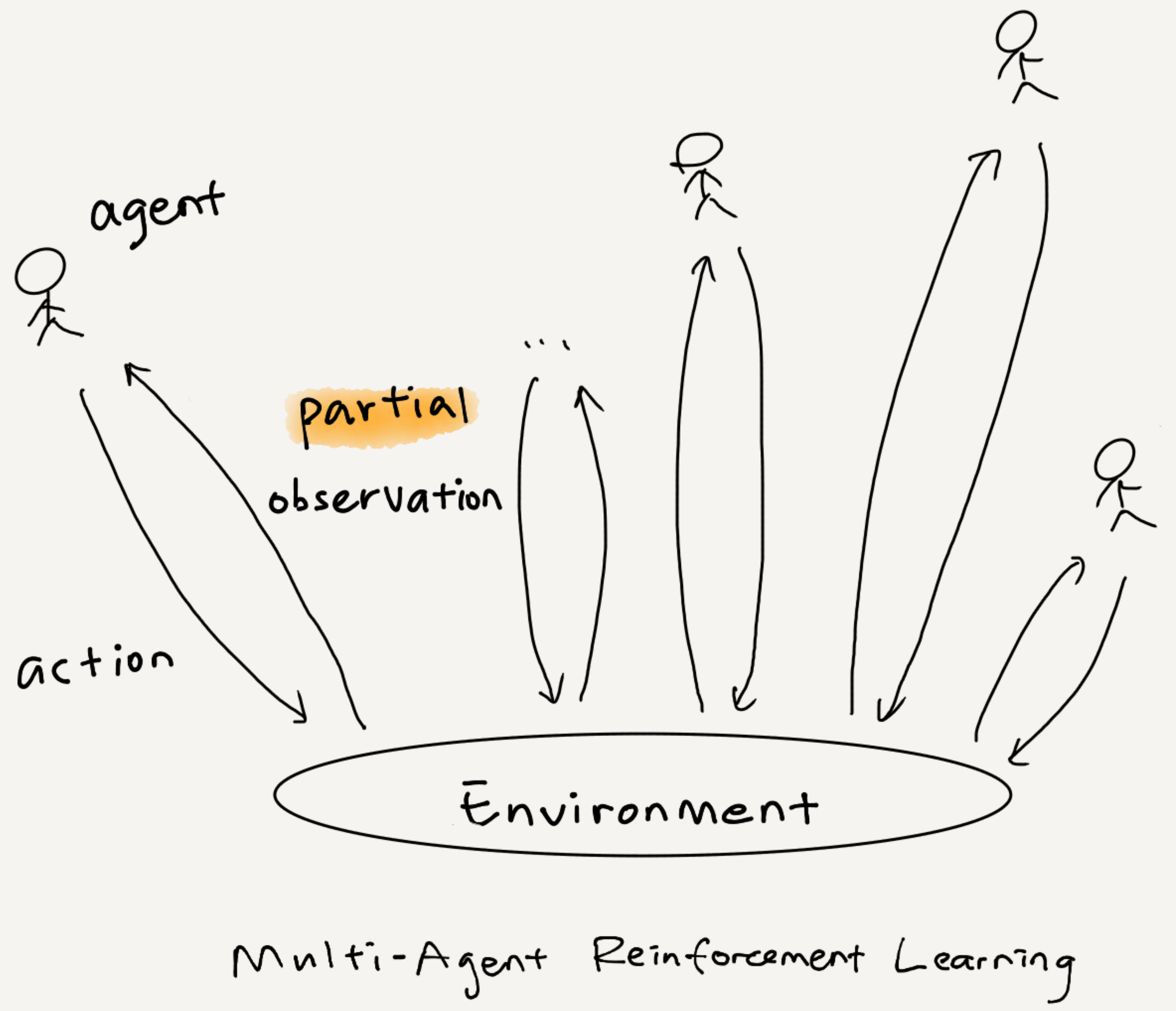


Multi-Agent Reinforcement Learning (MARL)



Multi-Agent Reinforcement Learning

Multi-Agent Reinforcement Learning (MARL)



Limitations of Existing Work

No communication for MARL-based edge caching

Zhong et al., “Deep multi-agent reinforcement learning based cooperative edge caching in wireless networks,” ICC 2019.
Wang et al., “Intelligent video caching at network edge: A multi-agent deep reinforcement learning approach,” INFOCOM 2020.

Limitations of Existing Work

No communication for MARL-based edge caching

- Partial visibility of the entire environment

Zhong et al., “Deep multi-agent reinforcement learning based cooperative edge caching in wireless networks,” ICC 2019.
Wang et al., “Intelligent video caching at network edge: A multi-agent deep reinforcement learning approach,” INFOCOM 2020.

Limitations of Existing Work

No communication for MARL-based edge caching

- Partial visibility of the entire environment

Few investigation of MARL communication in real-world applications

Zhong et al., "Deep multi-agent reinforcement learning based cooperative edge caching in wireless networks," ICC 2019.
Wang et al., "Intelligent video caching at network edge: A multi-agent deep reinforcement learning approach," INFOCOM 2020.

Limitations of Existing Work

No communication for MARL-based edge caching

- Partial visibility of the entire environment

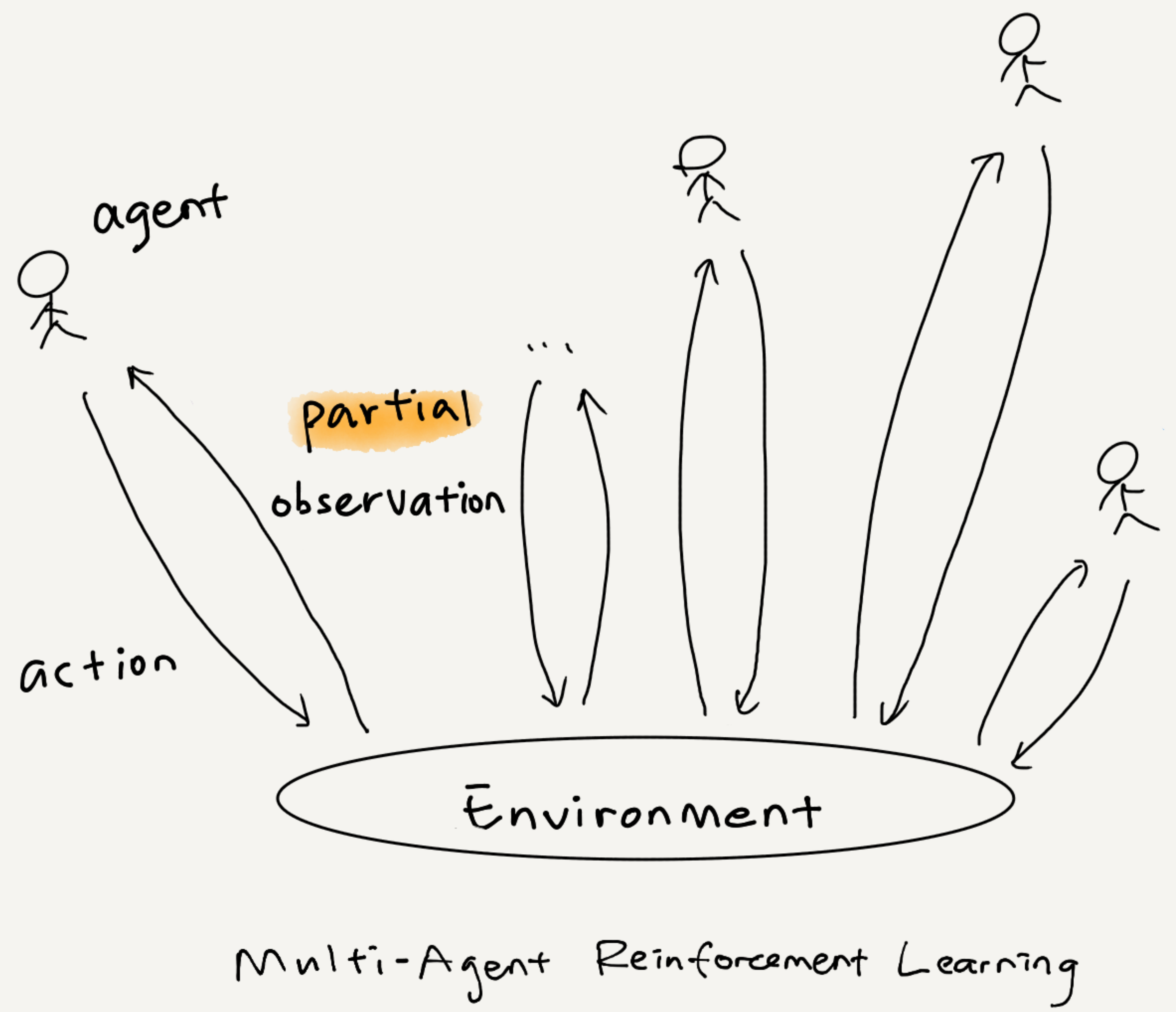
Few investigation of MARL communication in real-world applications

- Bandwidth constraints and delay

Zhong et al., “Deep multi-agent reinforcement learning based cooperative edge caching in wireless networks,” ICC 2019.
Wang et al., “Intelligent video caching at network edge: A multi-agent deep reinforcement learning approach,” INFOCOM 2020.

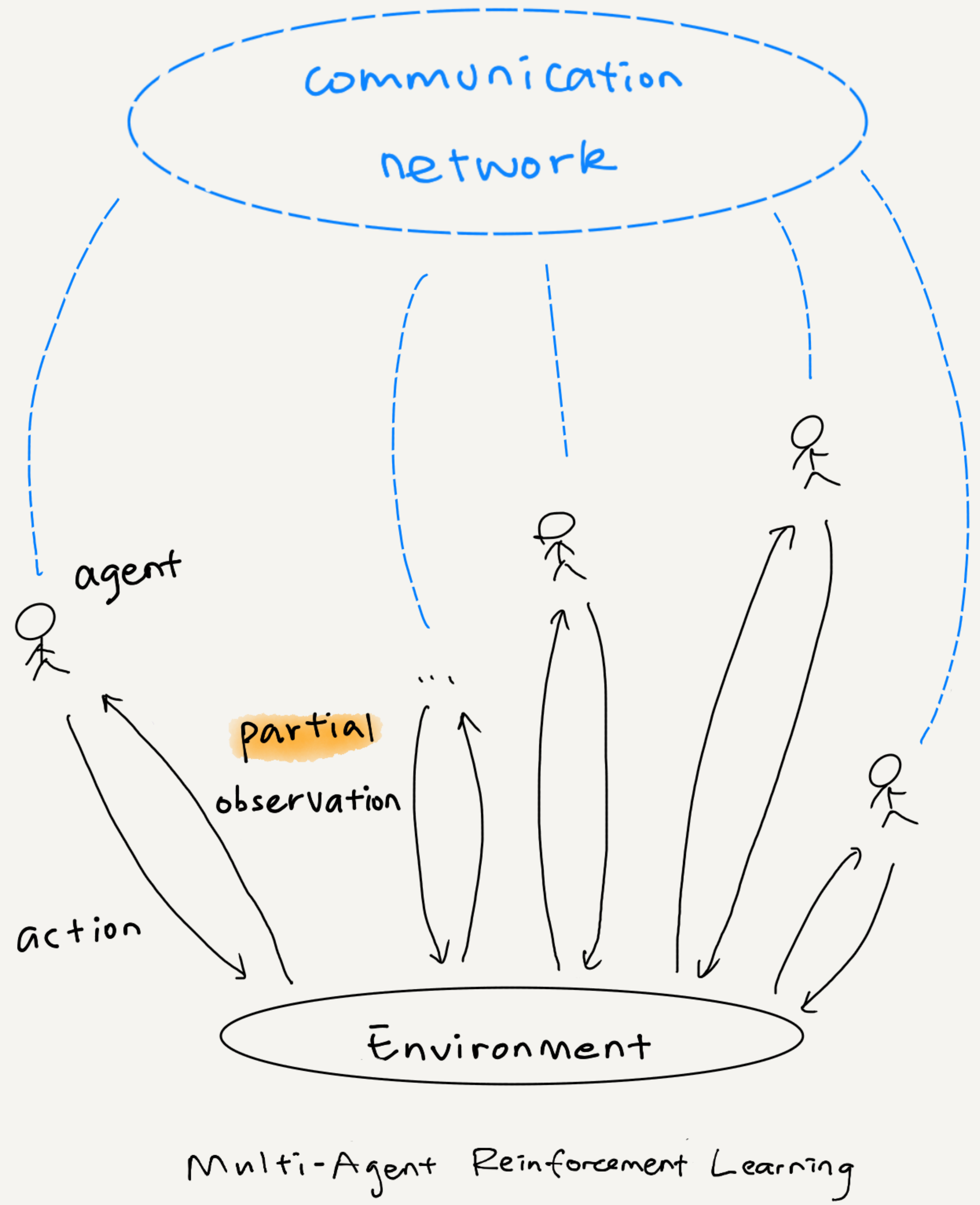
Objective

To design a practical multi-agent communication strategy that **facilitates MARL in real-world decision-making processes under bandwidth constraints**



Objective

To design a practical multi-agent communication strategy that **facilitates MARL in real-world decision-making processes under bandwidth constraints**



Learning to Communicate in Multi-Agent Cooperation

Intra-step communication[1-4]

Inter-step communication[5,6]

[1] Sukhbaatar et al., "Learning multiagent communication with backpropagation," NeurIPS 2016.
[2] Jiang et al., "Learning attentional communication for multi-agent cooperation," NeurIPS 2018.
[3] Zhang et al., "Efficient communication in multi-agent reinforcement learning via variance based control," NeurIPS 2019.
[4] Ding et al., "Learning individually inferred communication for multi-agent cooperation," NeurIPS 2020.
[5] Foerster et al., "Learning to communicate with deep multi-agent reinforcement learning," NeurIPS 2016.
[6] Das et al., "Tarmac: Targeted multi-agent communication," ICML 2019.

Learning to Communicate in Multi-Agent Cooperation

Intra-step communication[1-4]

- Introduces a considerable delay to the decision making process

Inter-step communication[5,6]

[1] Sukhbaatar et al., “Learning multiagent communication with backpropagation,” NeurIPS 2016.
[2] Jiang et al., “Learning attentional communication for multi-agent cooperation,” NeurIPS 2018.
[3] Zhang et al., “Efficient communication in multi-agent reinforcement learning via variance based control,” NeurIPS 2019.
[4] Ding et al., “Learning individually inferred communication for multi-agent cooperation,” NeurIPS 2020.
[5] Foerster et al., “Learning to communicate with deep multi-agent reinforcement learning,” NeurIPS 2016.
[6] Das et al., “Tarmac: Targeted multi-agent communication,” ICML 2019.

Learning to Communicate in Multi-Agent Cooperation

Intra-step communication[1-4]

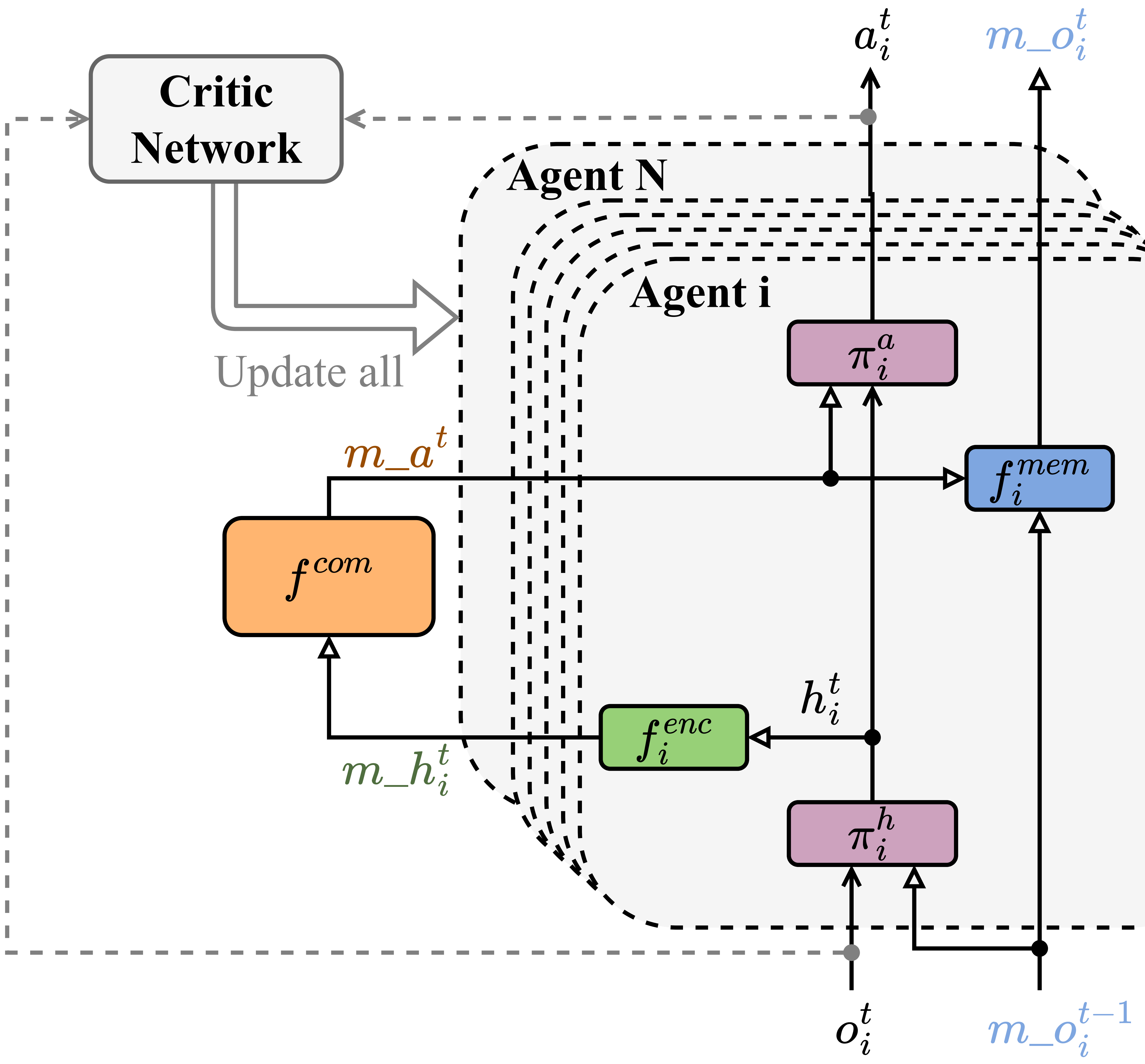
- Introduces a considerable delay to the decision making process

Inter-step communication[5,6]

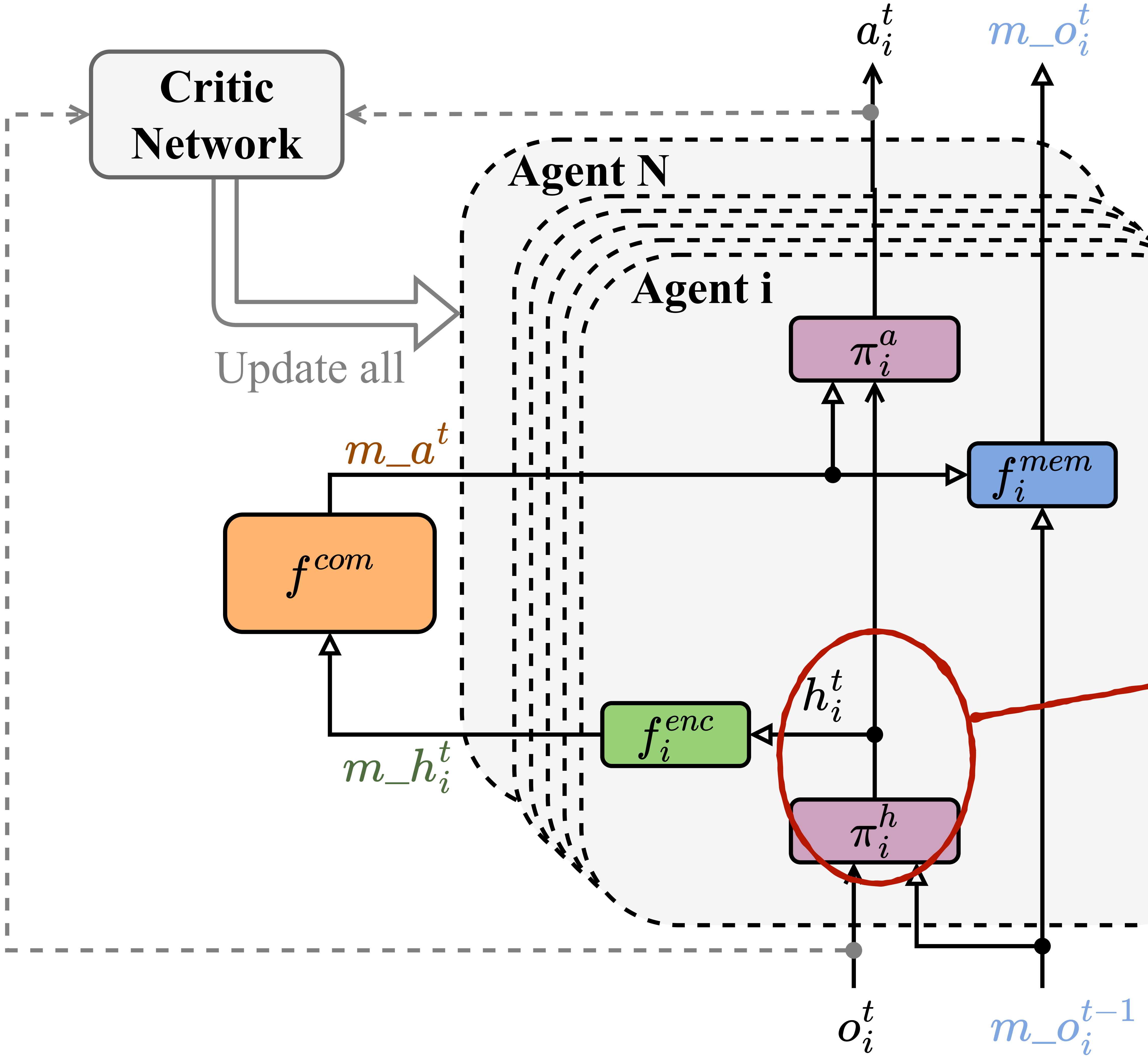
- Incorporates only the information or experience from the previous steps

[1] Sukhbaatar et al., "Learning multiagent communication with backpropagation," NeurIPS 2016.
[2] Jiang et al., "Learning attentional communication for multi-agent cooperation," NeurIPS 2018.
[3] Zhang et al., "Efficient communication in multi-agent reinforcement learning via variance based control," NeurIPS 2019.
[4] Ding et al., "Learning individually inferred communication for multi-agent cooperation," NeurIPS 2020.
[5] Foerster et al., "Learning to communicate with deep multi-agent reinforcement learning," NeurIPS 2016.
[6] Das et al., "Tarmac: Targeted multi-agent communication," ICML 2019.

Our Design: Hybrid Communication

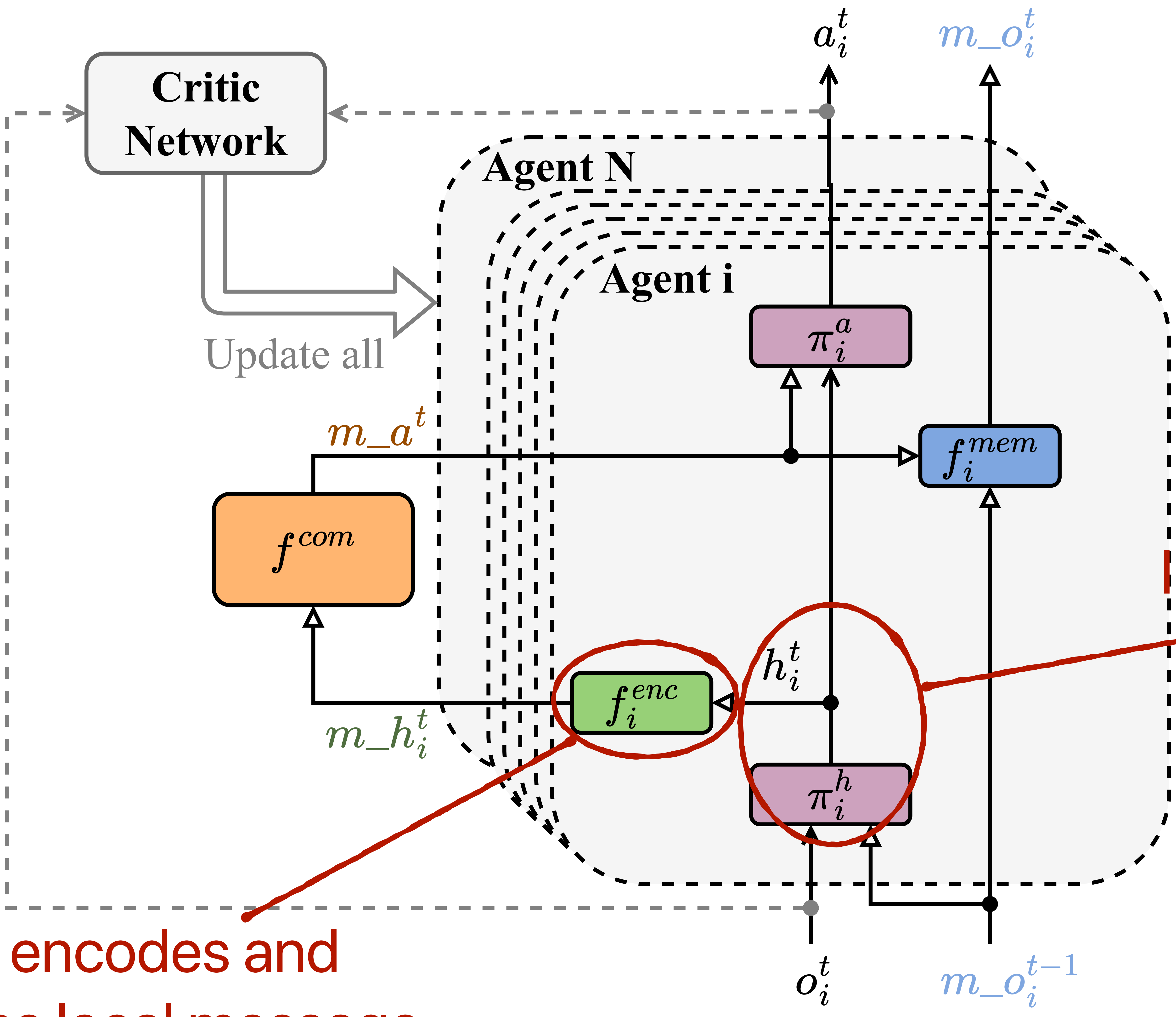


Our Design: Hybrid Communication



Incorporates knowledge about observation and action intention into local message

Our Design: Hybrid Communication

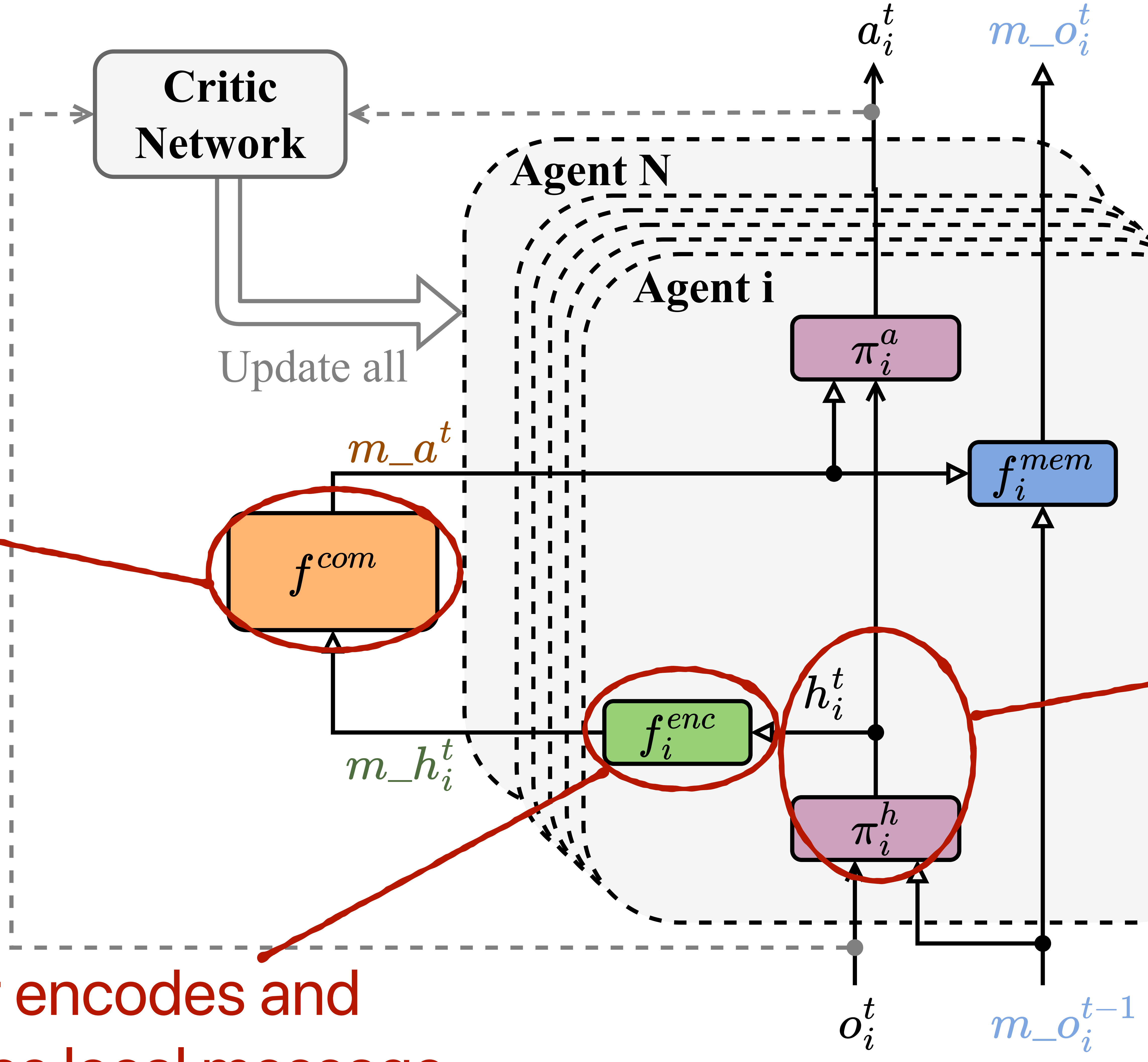


Incorporates knowledge about observation and action intention into local message

Further encodes and compresses local message

Our Design: Hybrid Communication

Broadcasts global message

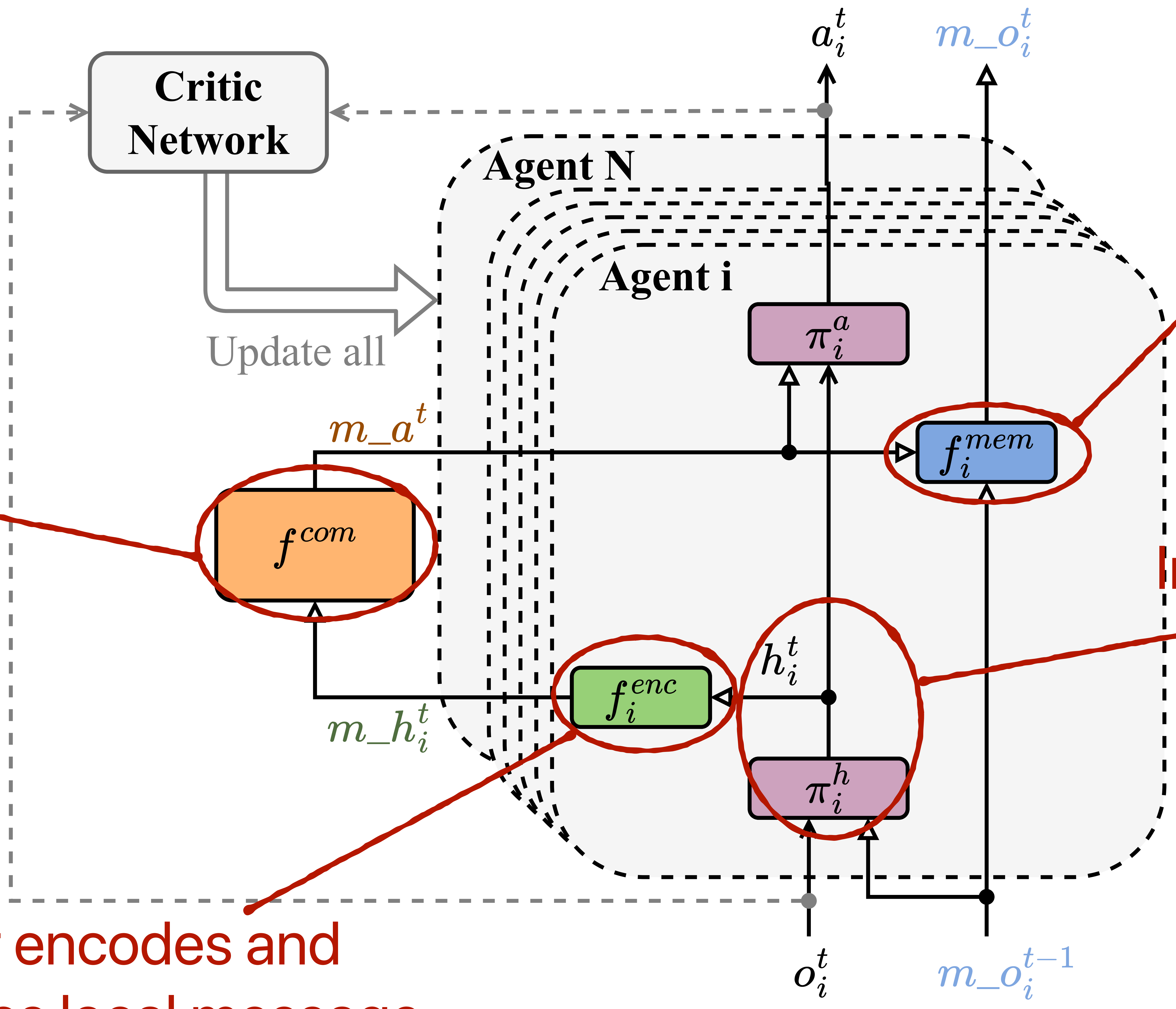


Incorporates knowledge about observation and action intention into local message

Further encodes and compresses local message

Our Design: Hybrid Communication

Broadcasts global message

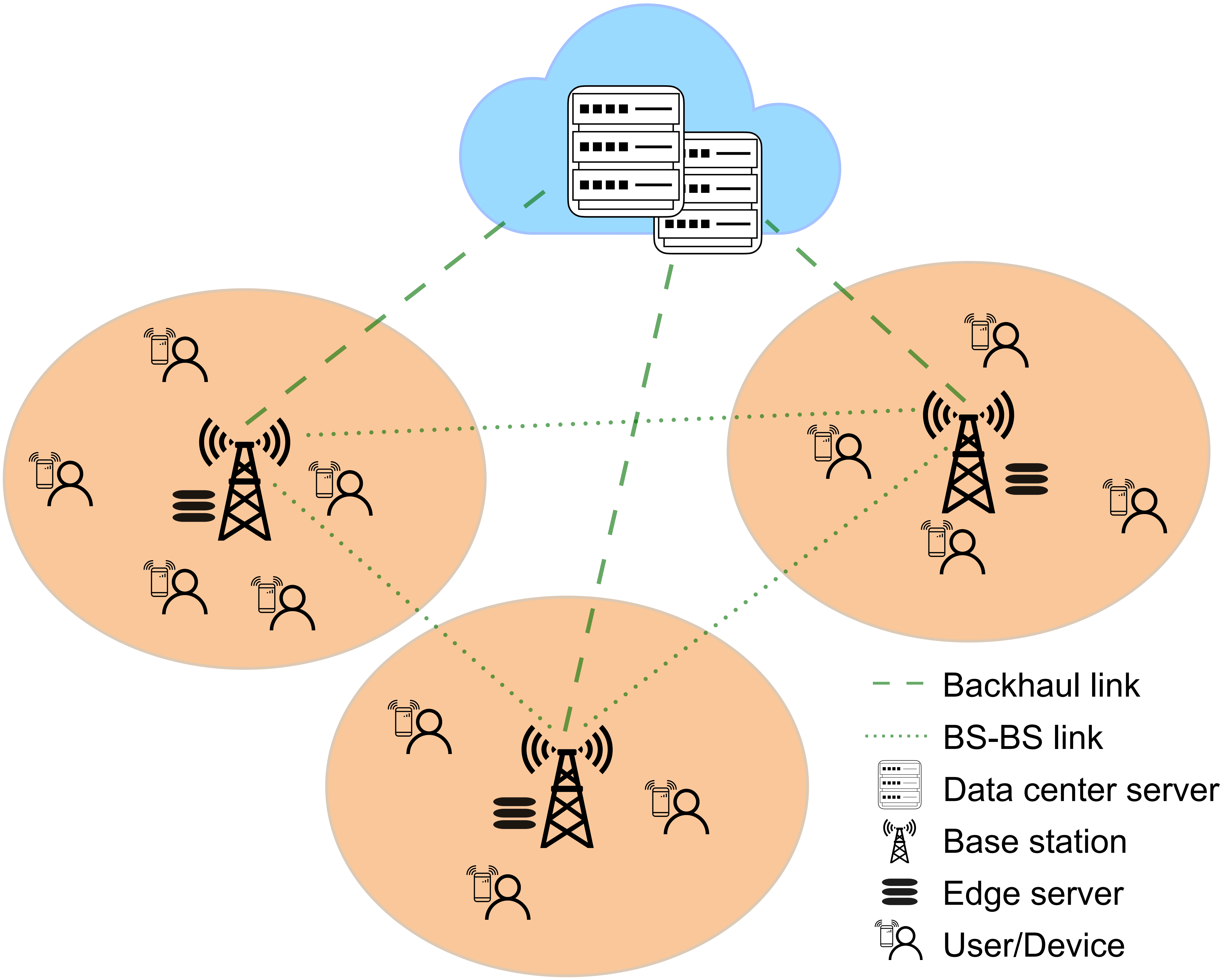


Incorporates both history and current knowledge

Incorporates knowledge about observation and action intention into local message

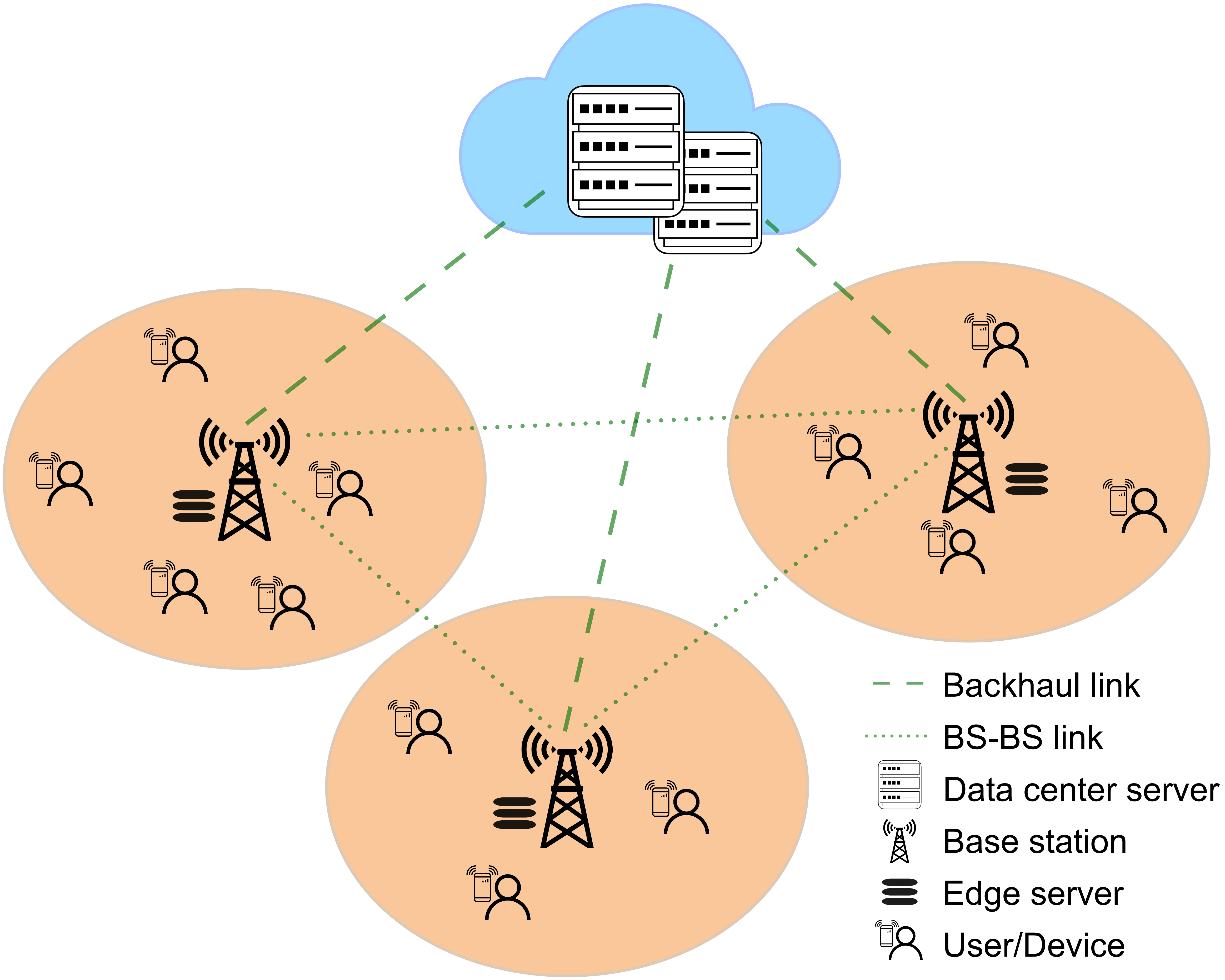
Further encodes and compresses local message

MARL Formulation for Edge Caching



MARL Formulation for Edge Caching

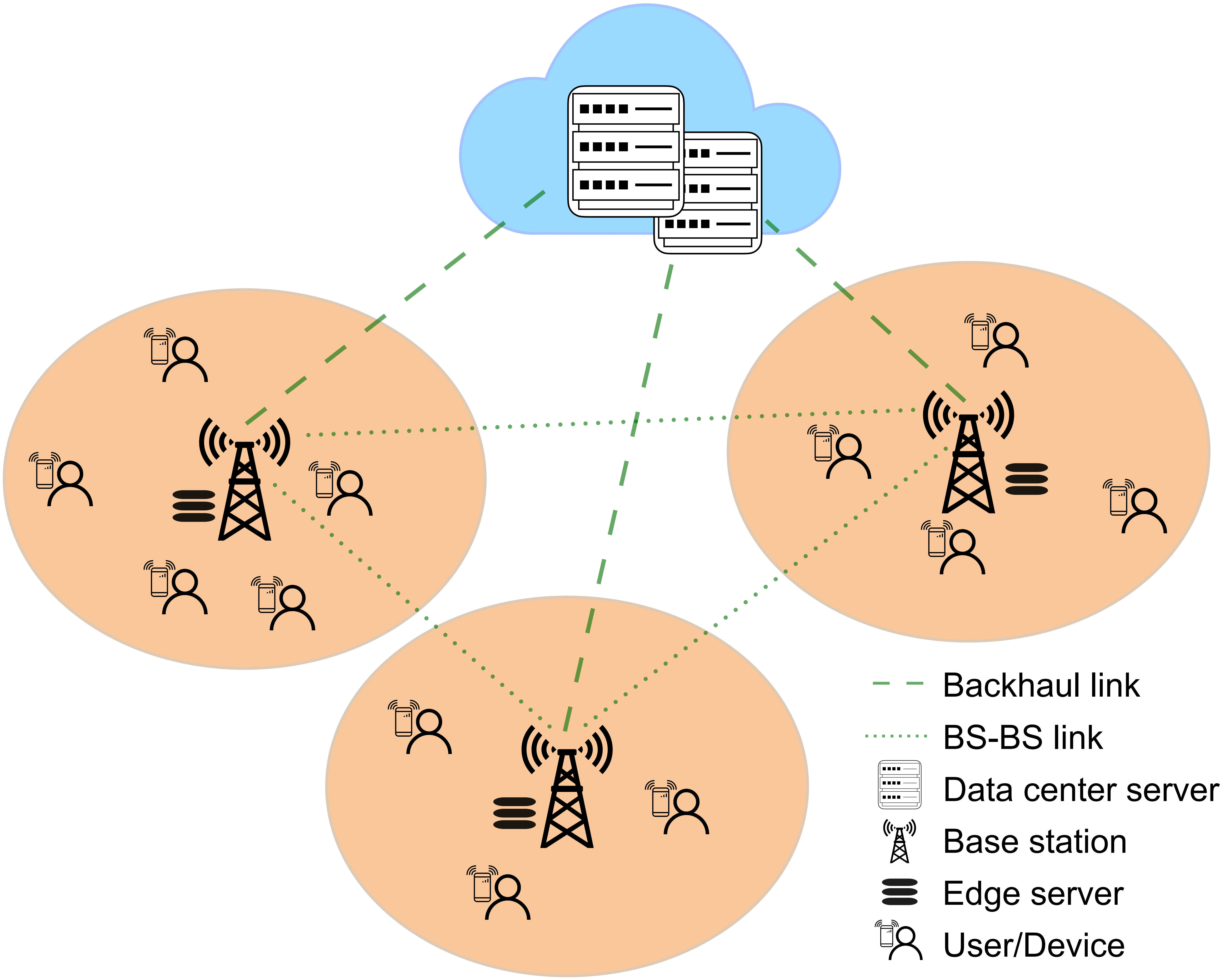
State



MARL Formulation for Edge Caching

State

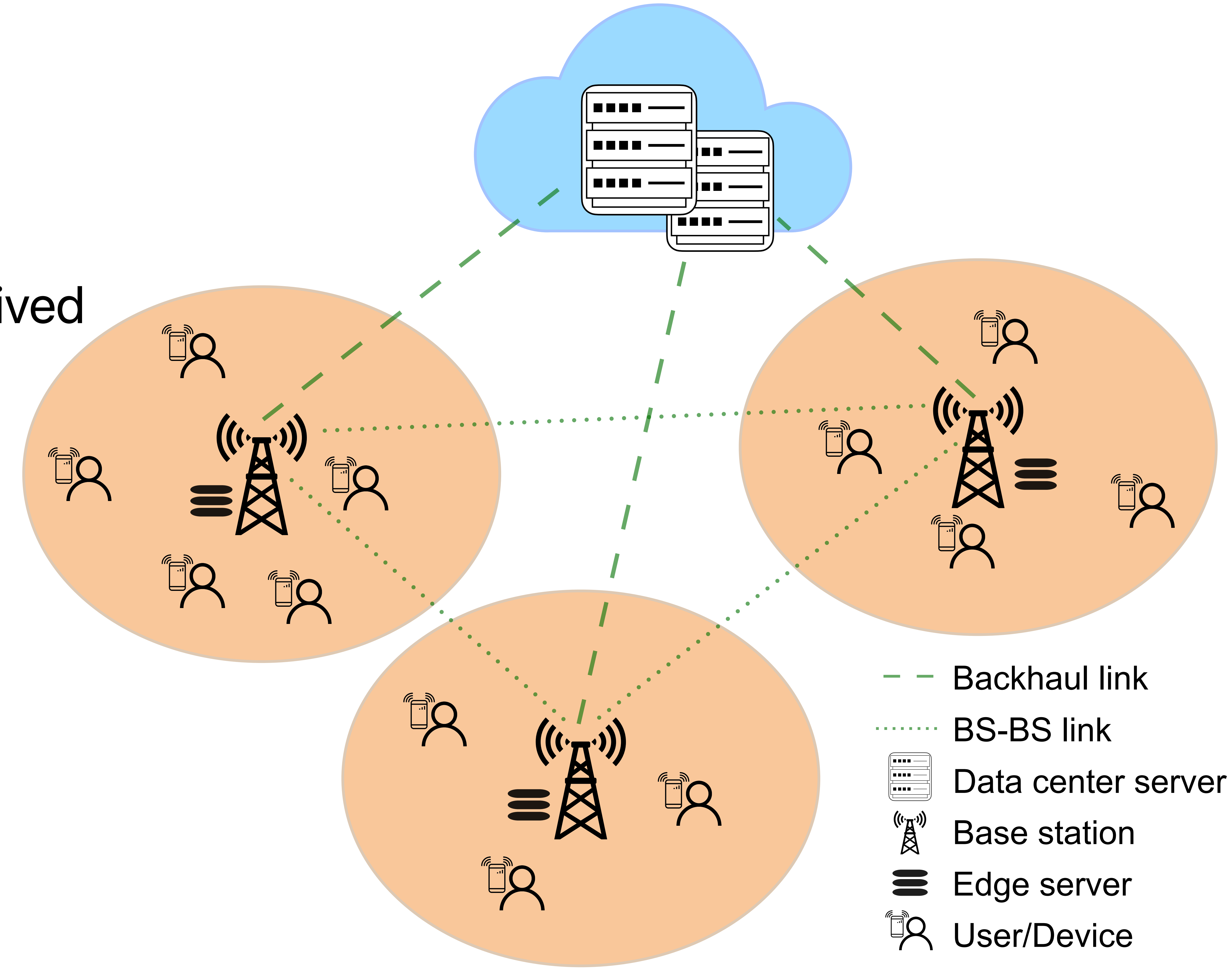
- Caching state: represents whether each content is cached locally or not



MARL Formulation for Edge Caching

State

- Caching state: represents whether each content is cached locally or not
- Request state: indicates the requested content received locally

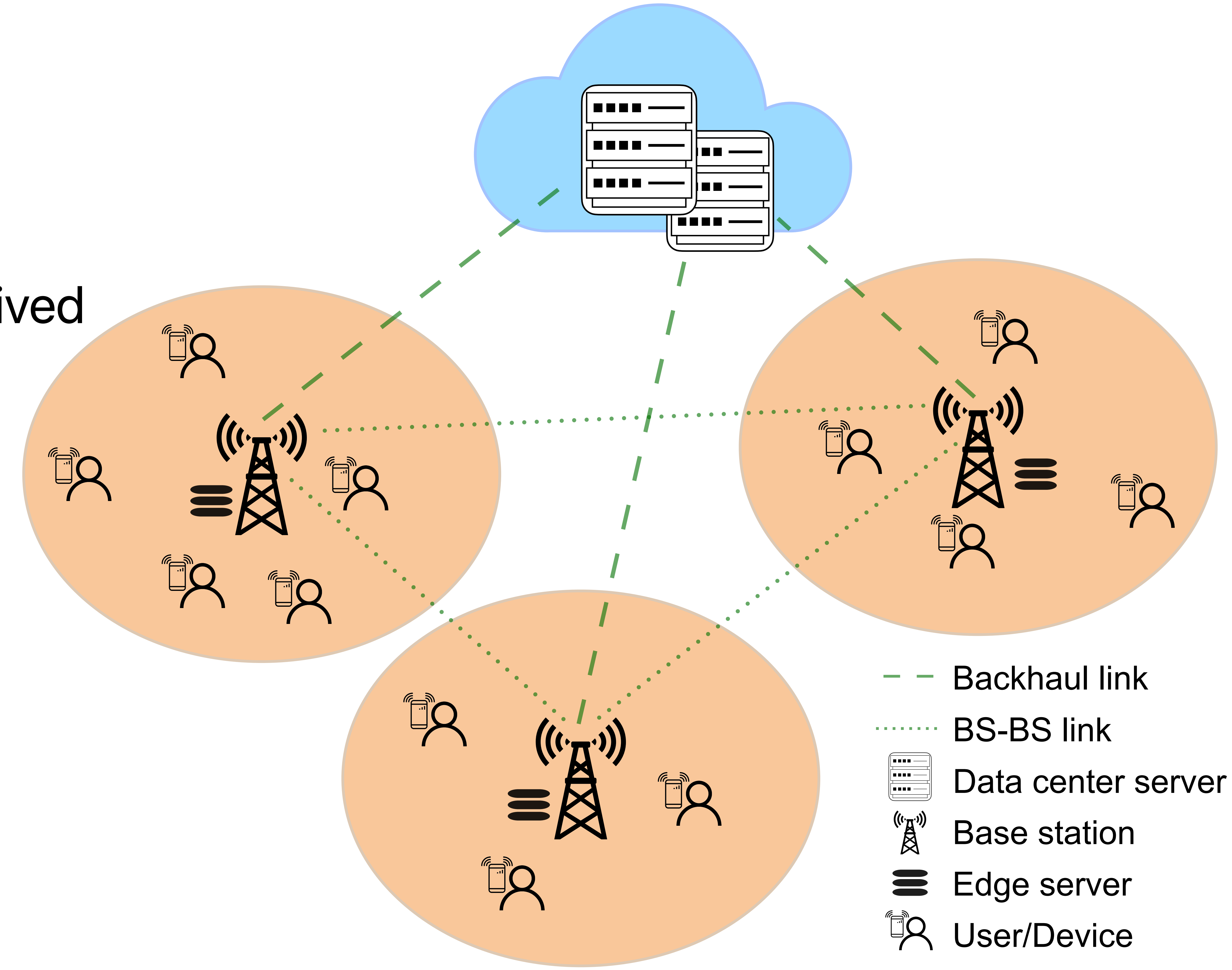


MARL Formulation for Edge Caching

State

- Caching state: represents whether each content is cached locally or not
- Request state: indicates the requested content received locally

Action



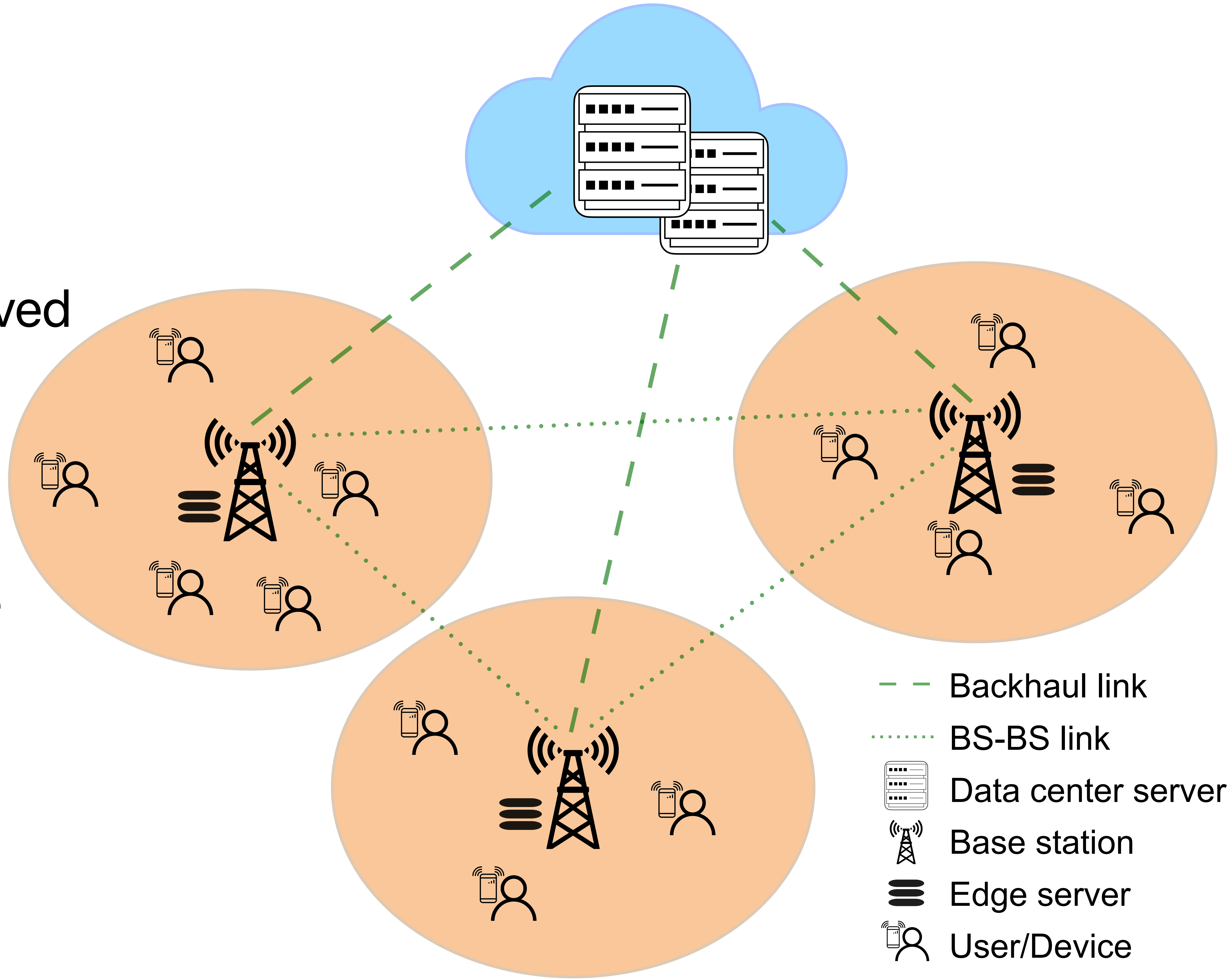
MARL Formulation for Edge Caching

State

- Caching state: represents whether each content is cached locally or not
- Request state: indicates the requested content received locally

Action

- Determines how important each content needs to be cached



MARL Formulation for Edge Caching

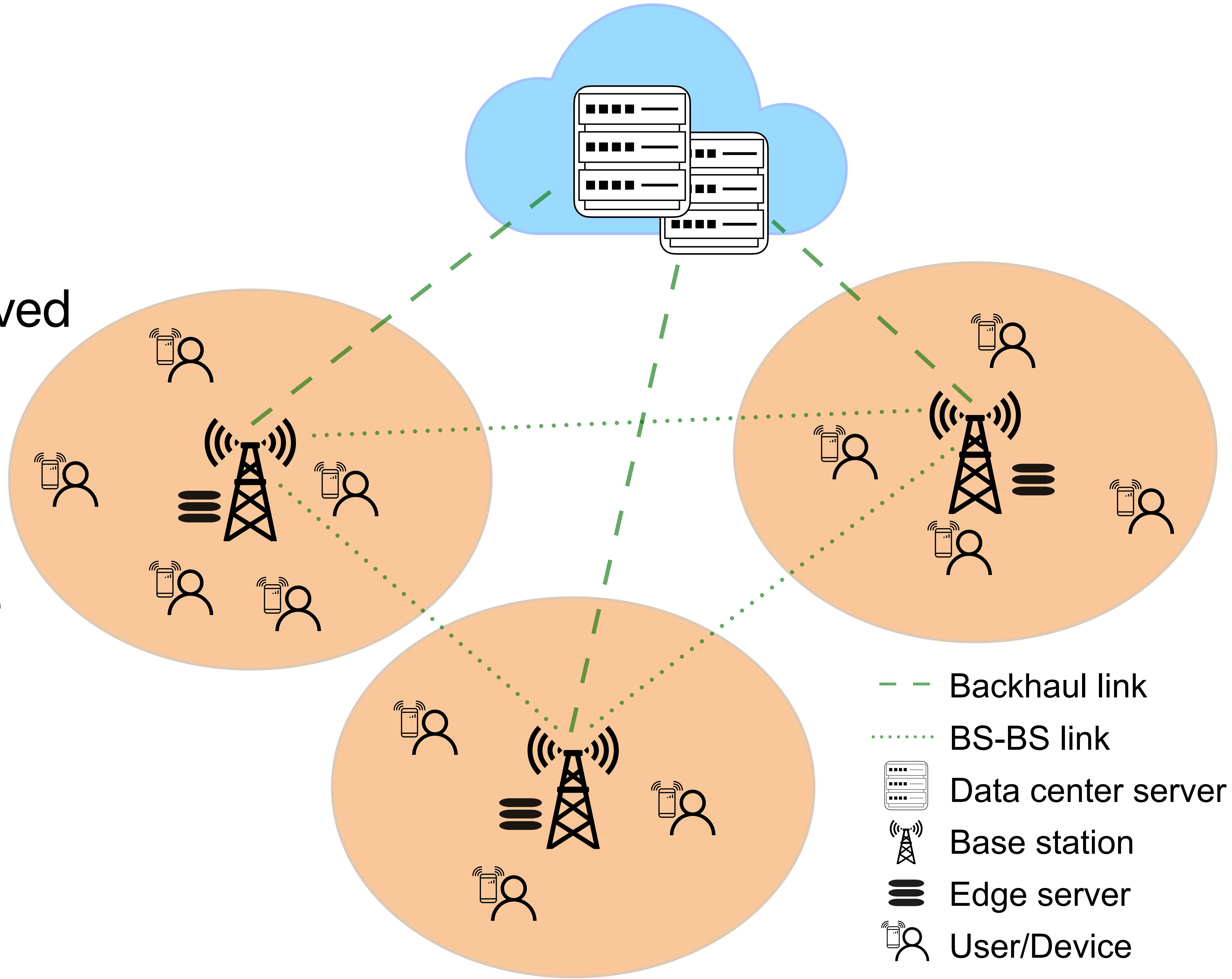
State

- Caching state: represents whether each content is cached locally or not
- Request state: indicates the requested content received locally

Action

- Determines how important each content needs to be cached

Reward



MARL Formulation for Edge Caching

State

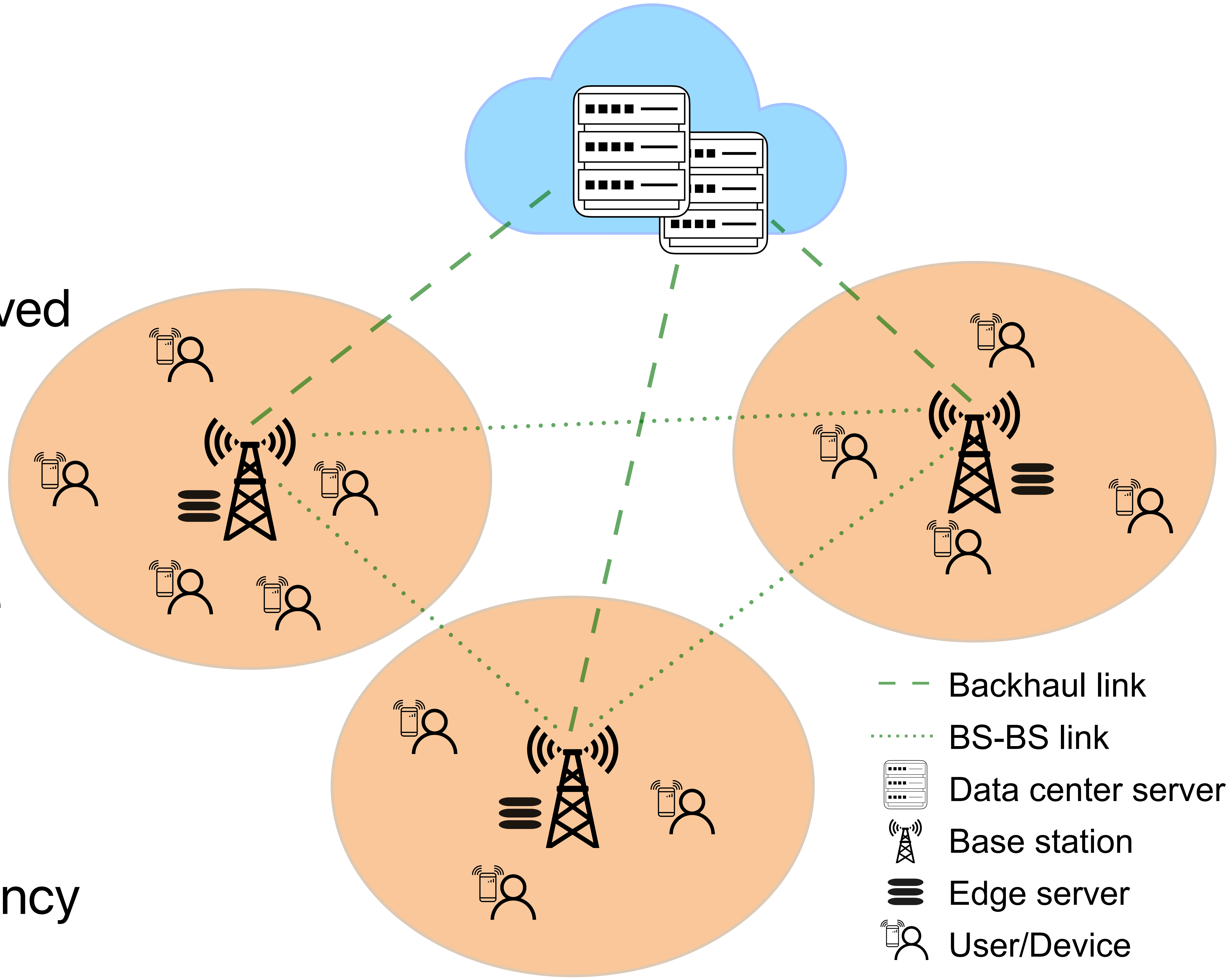
- Caching state: represents whether each content is cached locally or not
- Request state: indicates the requested content received locally

Action

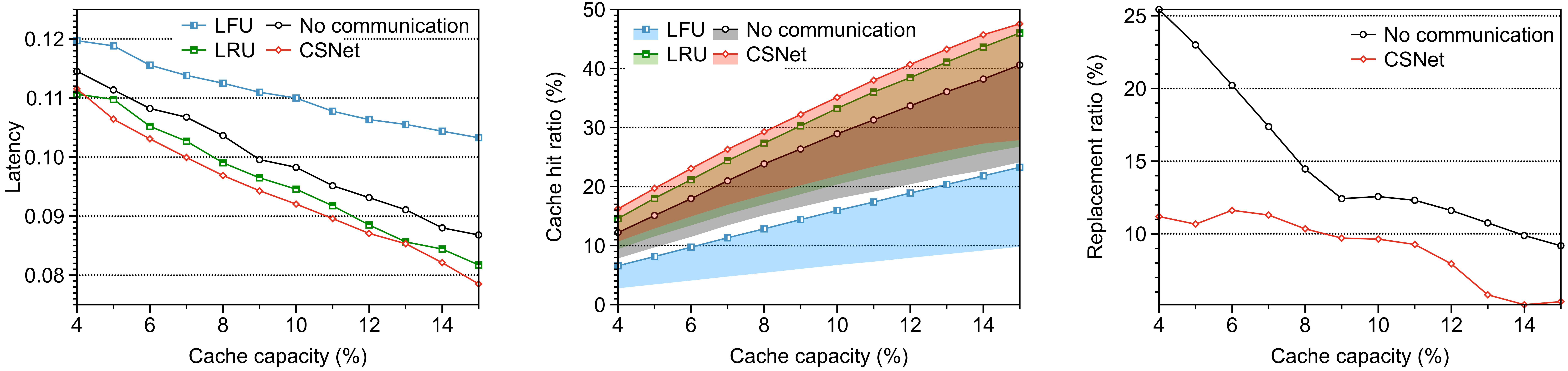
- Determines how important each content needs to be cached

Reward

- Credit for cache hit rate, penalty for transmission latency and cache replacement cost



Caching Performance



(a) Latency (b) Total (local + neighbor) hit ratio (c) Replacement ratio

Fig. 3: The latency, cache hit ratio, replacement ratio of different policies under varying cache capacities.

Caching Performance

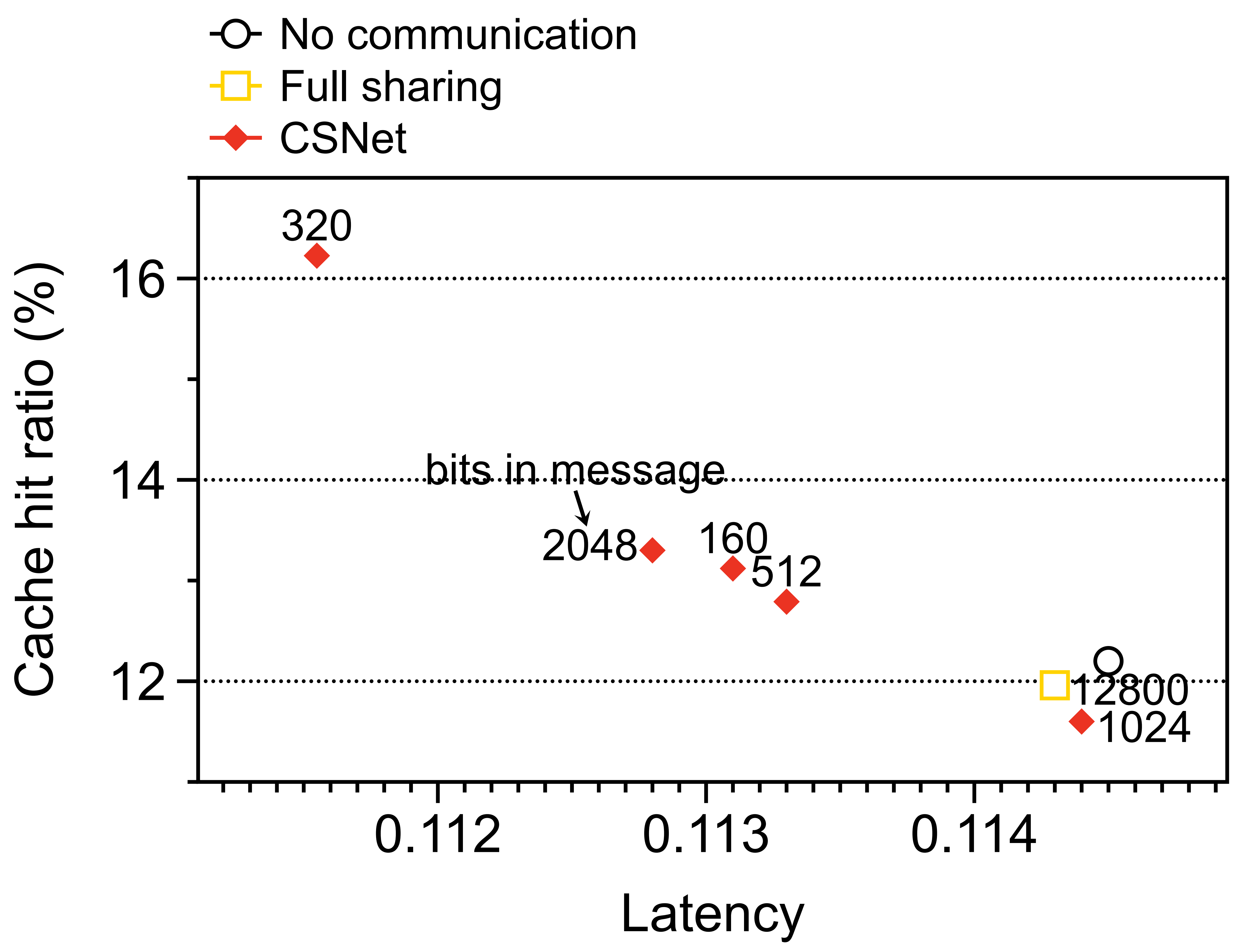


Fig. 4: The average hit ratio and latency of different message size under 4% cache capacity.

Conclusion

Conclusion

Our approach closes the gap between **learning-based caching mechanism** and **learning-based multi-agent communication**

Conclusion

Our approach closes the gap between **learning-based caching mechanism** and **learning-based multi-agent communication**

- Outperforms other caching algorithms that are rule-based, MARL-based without communication, or MARL-based with full observations sharing

Conclusion

Our approach closes the gap between **learning-based caching mechanism** and **learning-based multi-agent communication**

- Outperforms other caching algorithms that are rule-based, MARL-based without communication, or MARL-based with full observations sharing

- Introduces limited communication overhead and delay considering bandwidth constraints

Thank you!