USF UNIVERSITY OF SOUTH FLORIDA
COLLEGE OF ENGINEERING
Department of Electrical Engineering

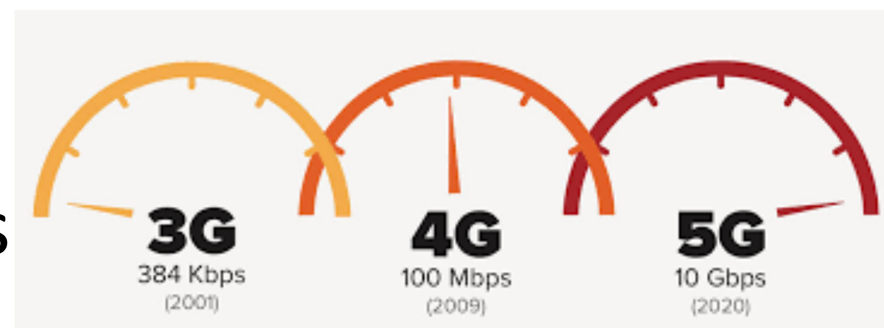
Communications for Artificial Intelligence

Professor Kwang-Cheng Chen, IEEE Fellow
Department of Electrical Engineering
University of South Florida
email: kwangcheng@usf.edu

Appreciations to research support from Cyber Florida

5G Technology

- The well known pillar technologies for 5G
  - Enhanced Mobile Broadband (eMBB)
  - Massive M2M/IoT Communication (mMTC)
  - Ultra Reliable and Low Latency Communication (uRLLC)
    - Mission critical services
- Field trials are coming
- Machine learning emerges
- AI for 5G and Beyond



Reality in 5G Technology

- Technology for eMBB is mature in 5G (R15)
  - C-RAN, massive MIMO, …
  - NOMA attracts extensive research but is not used
- mMTC is getting mature, after years' efforts

RECENT PROGRESS IN MACHINE-TO-MACHINE COMMUNICATIONS

IEEE Comm. Mag April 2011

Toward Ubiquitous Massive Accesses in 3GPP Machine-to-Machine Communications

Shao-Yu Lien and Kwang-Cheng Chen, NTU
Yonghua Lin, IBM Research Division

- uRLLC is rather young and only one use case in R15
  - It is actually technological paradigm shift, but not widely understood
  - Reliability suggests successful completion of intelligent control and management missions, not just successful packet delivery
  - A packet correctly received but exceeding latency requirement is useless
  - Primary M2M, not H2H

Human Intelligence

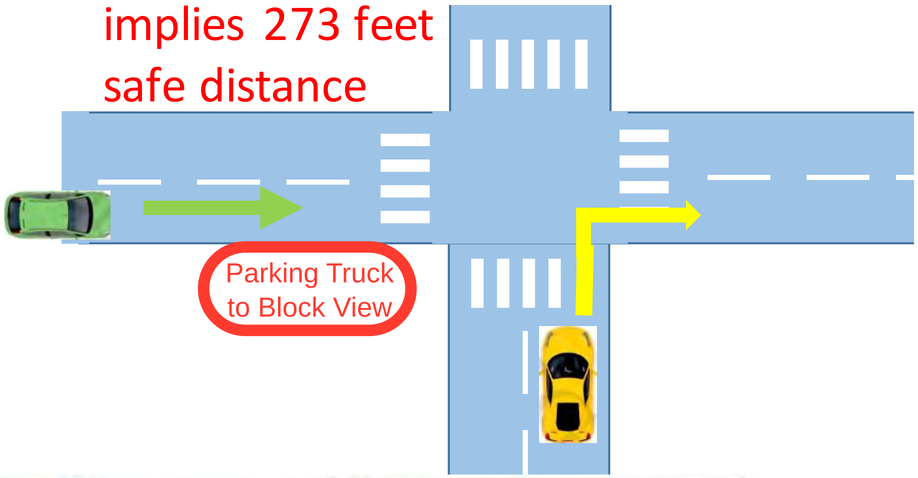
Are you still making phone calls?

AI, AI, and AI; ML, ML, and ML

RISE OF MACHINES

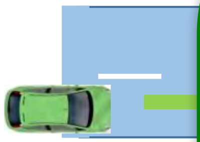
From Driving Assistance to Autonomous Driving

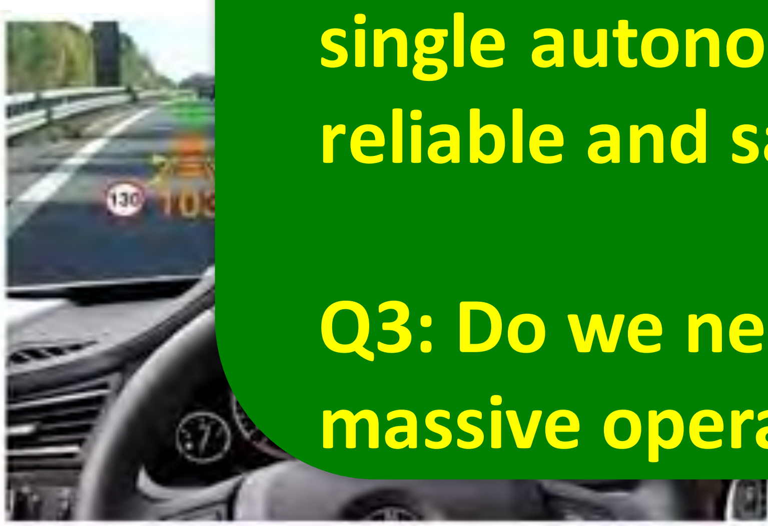
Green car at 50 MPH implies 273 feet safe distance



UNIVERSITY OF SOUTH FLORIDA
COLLEGE OF ENGINEERING
Department of Electrical Engineering

From Driving Assistance to Autonomous Driving





Q1: Do we need networking for a single autonomous vehicle as good as human driving?

Q2: Do we need networking for a single autonomous vehicle more reliable and safe than human driving?

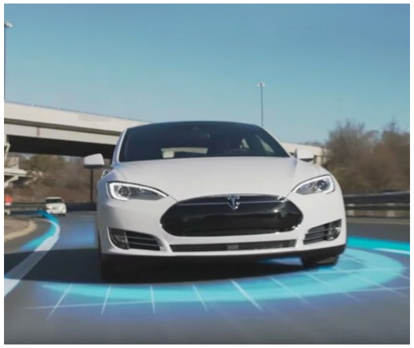
Q3: Do we need networking for massive operation of AVs?

An Illustration by Autonomous Vehicles on Manhattan Streets [IEEE Globecom 2018]

WIRELESS COMMUNICATIONS MEETS ARTIFICIAL INTELLIGENCE

Autonomous Vehicles and Mobile Riobots

- Autonomous vehicles are increasingly present in modern society
- The development of AVs is the key of the new transportation system

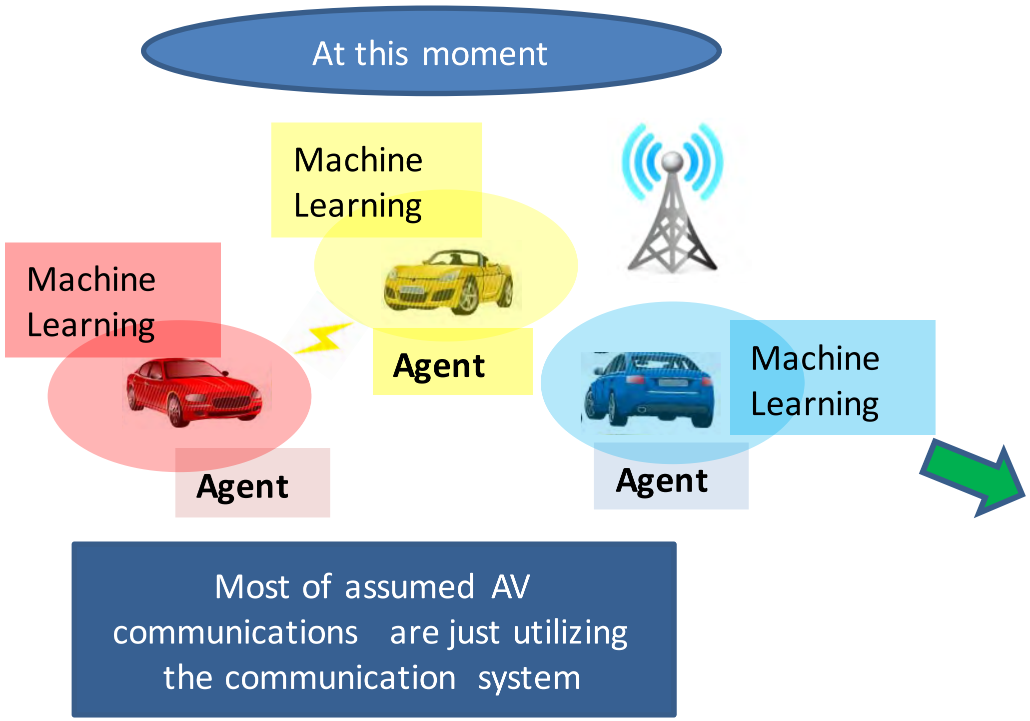


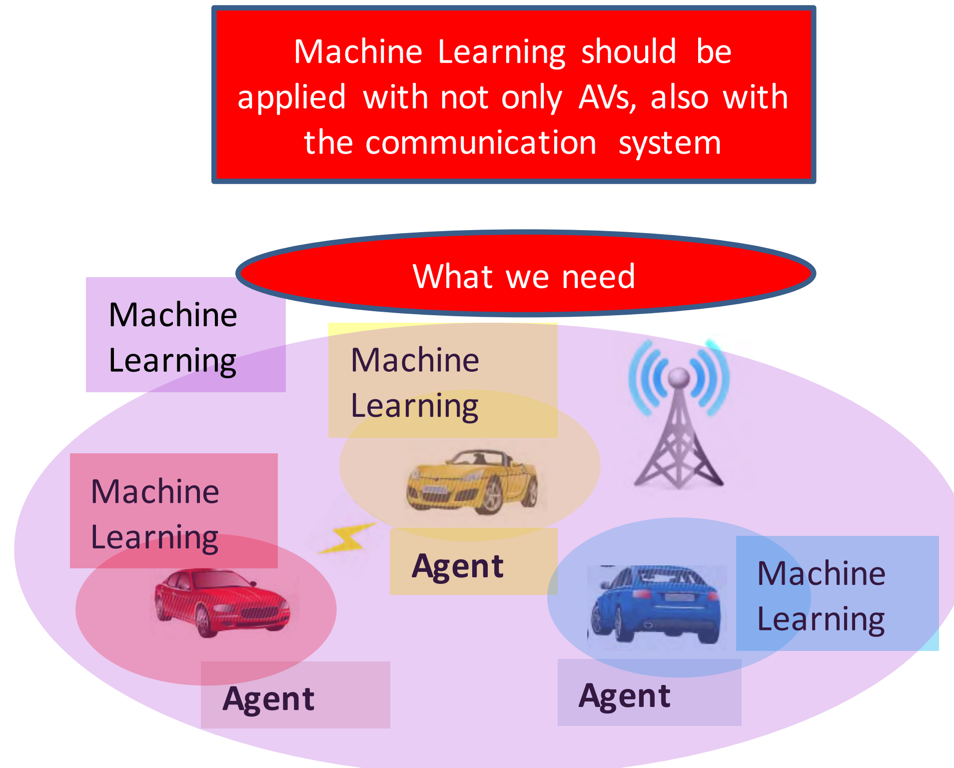
- The technologies of Autonomous Vehicle itself are developed well
- But we assume a lot of AVs on the road
- Communications among AVs will be crucial

Image : govtech.com

UNIVERSITY OF SOUTH FLORIDA
COLLEGE OF ENGINEERING
Department of Electrical Engineering

Technological Requirements



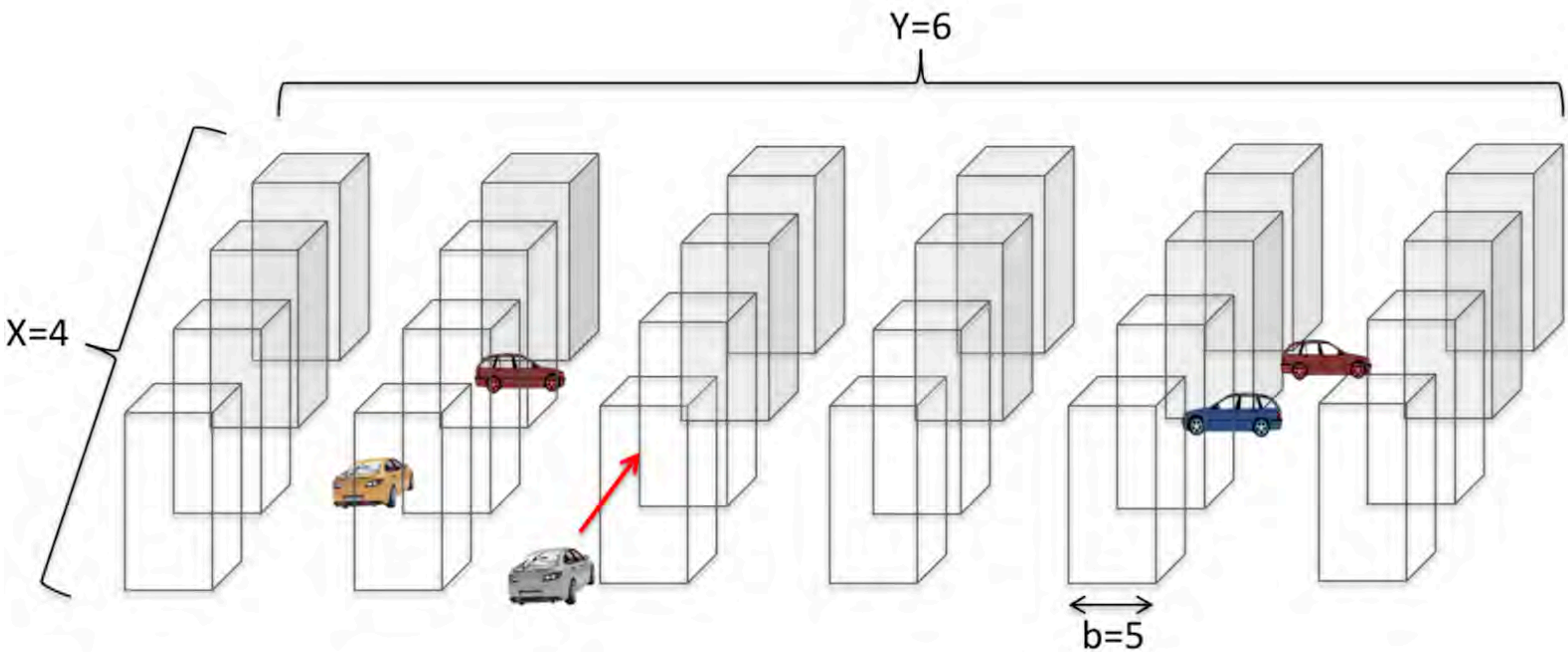


- Radio resource for bandwidth or latency?
- What to communication?
- How and when to communicate?

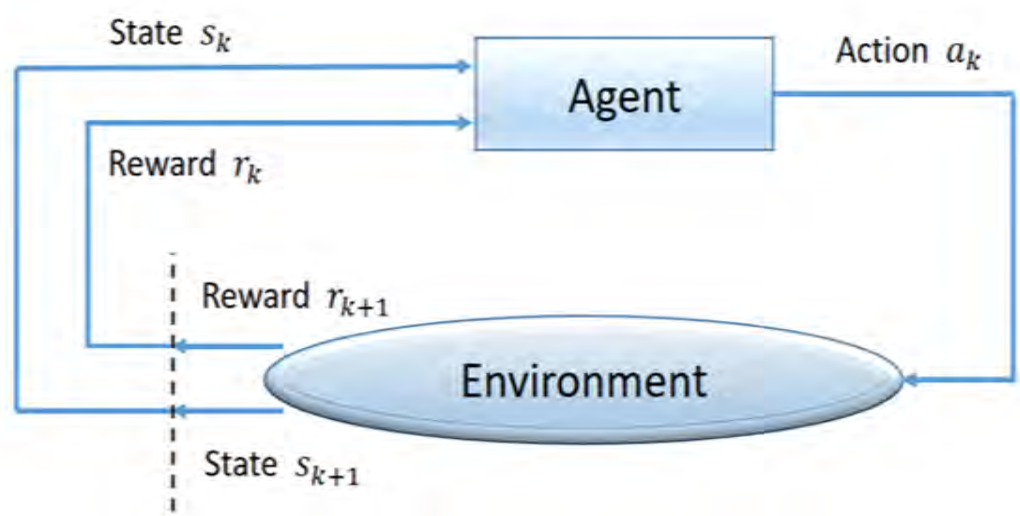
Deeper Technological Thoughts

- If AI Agents (autonomous vehicles, robots, etc.) in massive scale, will be around everywhere in our daily life, their interactions would be the most critical issue in future digital society.
- The role of (wireless) networking and communication for AI/ML is overlooked in literature, we know almost NOTHING from this aspect.
  - We assume M2M communication identical to H2H communication. **Is it right?**

A Toy Model for AVs over Manhattan Streets



Reinforcement Learning



- The agent implements a mapping from states to probabilities of selecting each possible action
- This mapping is called the agent's policy and is denoted π_k, where $\pi_k(a|s)$ is the probability that $a_k = \text{a}$ if $s_k = s$
- **The agent's goal is to maximize the total amount of reward it receives over the long run**

Value Function

Value Function

- Functions of states (or of state-action pairs) that estimate **how good** it is for the agent to be in a given state (or how good it is to perform a given action in a given state)
- **A policy**, π, is a mapping from each state, $s \in \mathcal{S}$, and action, $a \in \mathcal{A}(s)$, to the probability $\pi(a|s)$ of taking action a when in state s
- **The value** of a state s under a policy π, denoted $v_\pi(s)$, is the expected return when starting in s and following π thereafter

$$v_\pi(s) = \mathbb{E}_\pi[G_k|s_k = s] = \mathbb{E}_\pi\left[\sum_{d=0}^{\infty} \gamma^d R_{k+d+1} \middle| s_k = s\right]$$

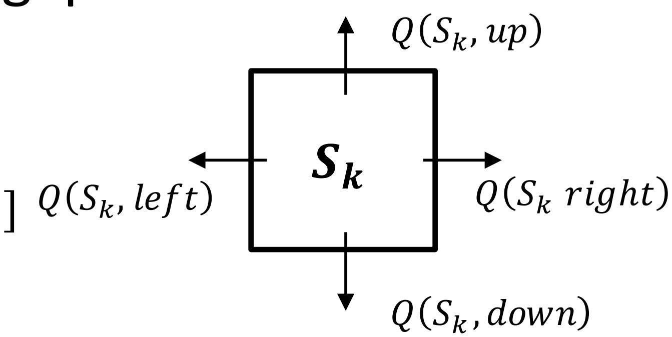
$\mathbb{E}_\pi[\cdot]$ denote the expected value of a random variable given that the agent follows policy π

Q-learning

- When state estimation is not perfectly available, using belief-action in RL, Q-learning.
- Under the environment changing, Q-learning is updating q-function $Q(s_k, a_k)$ with q-function $Q(s_{k+1}, a_k)$ at next state
- α is learning step, which is related to learning speed

$$Q(s_k, a_k) \leftarrow (1-\alpha)Q(s_k, a_k) + \alpha[r_{k+1} + \gamma \max_a Q(s_{k+1}, a)]$$

$$Q(s_k, a_k) \leftarrow Q(s_k, a_k) + \alpha[r_{k+1} + \gamma \max_a Q(s_{k+1}, a) - Q(s_k, a_k)]$$

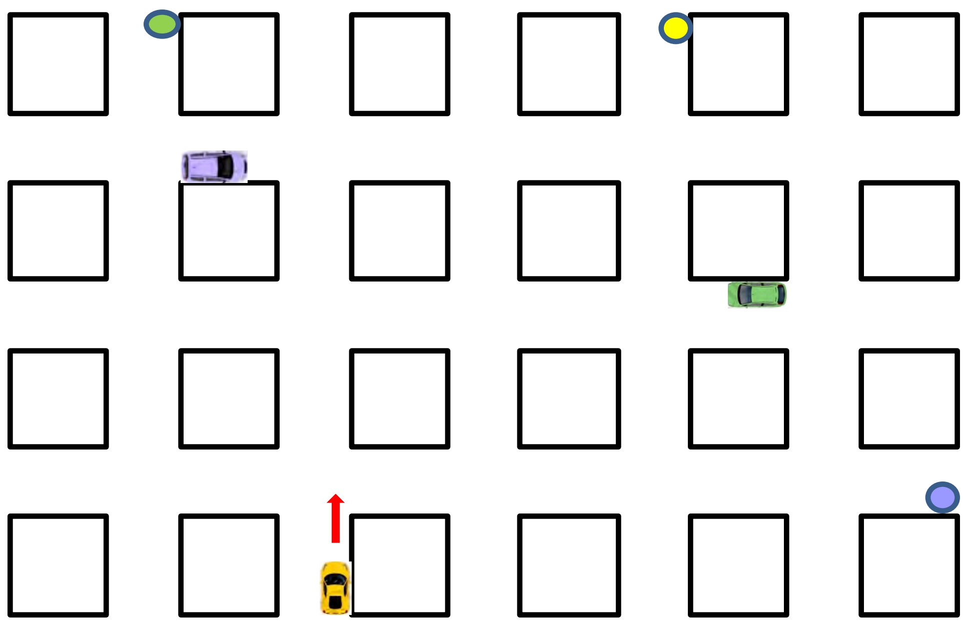


What we need to consider

- The length of planning horizon
- The reward structure

Manhattan Street Model

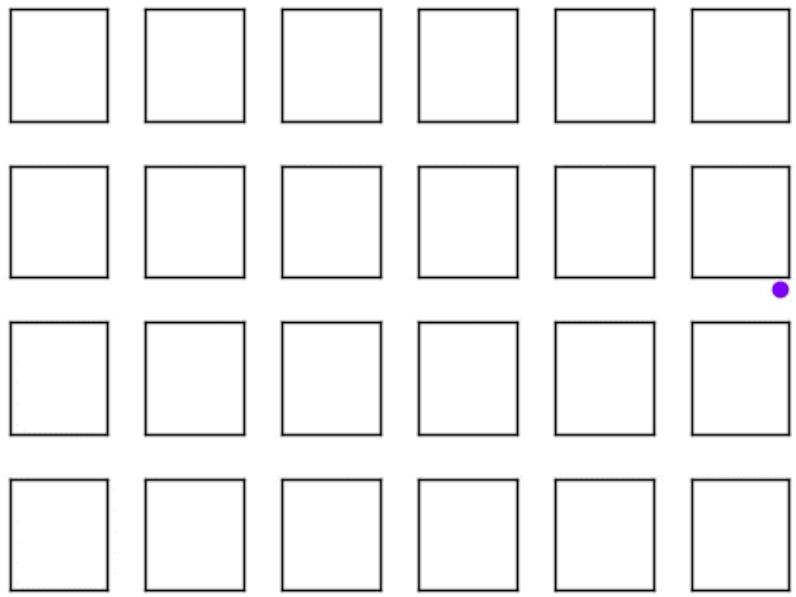
- Each vehicle has different source-destination relations but going through the same street region without knowing each others' information
- Public reference (i.e. map) is available to all vehicles



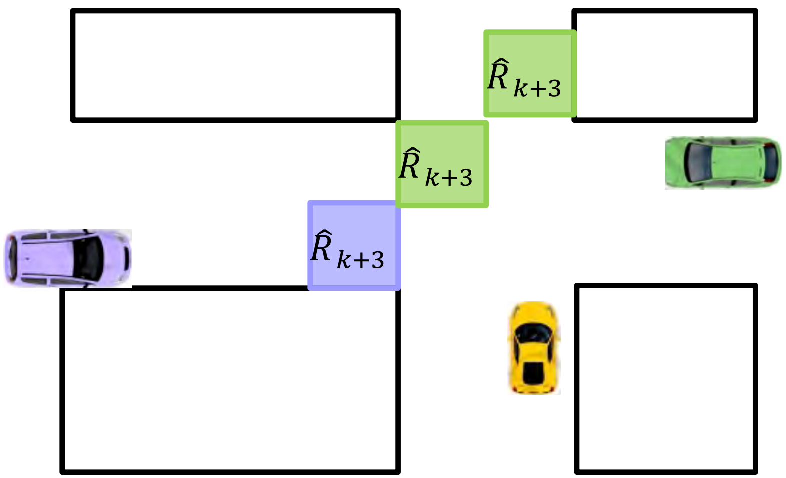
USF UNIVERSITY OF SOUTH FLORIDA
COLLEGE OF ENGINEERING
Department of Electrical Engineering

Manhattan Street Demo

- Manhattan Model Street (M=4, N=6, b=5)
- 10 cars coming into the street
- Safety and reliability is top priority over time to travel (performance)



Without Communication



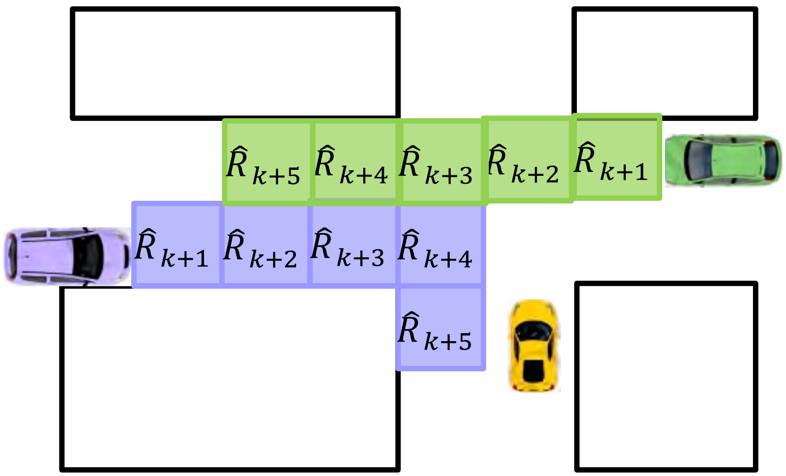
- At time k, the agent recognizes other vehicle and generates reward map $R_{i,k}$
- For $k+d$ $(d=1,\ldots,D, D$ is the depth of horizon) will be the expected reward map $\hat{R}_{k+d}$

$$\hat{R}_{i,k+d} = [\hat{r}_{s_{k+d}}]$$

$$\mathbb{R}_{i,k:k+D} = \{R_{i,k}, \hat{R}_{i,k+1}, \ldots, \hat{R}_{k+D}\}$$

The agent makes decision based on $\mathbb{R}_{i,k:k+D}$

Without Communication



- At time k, the agent recognizes other vehicle and generates reward map $R_{i,k}$
- For $k+d$ $(d = 1, \ldots, D, D$ is the depth of horizon) will be the expected reward map $\hat{R}_{k+d}$

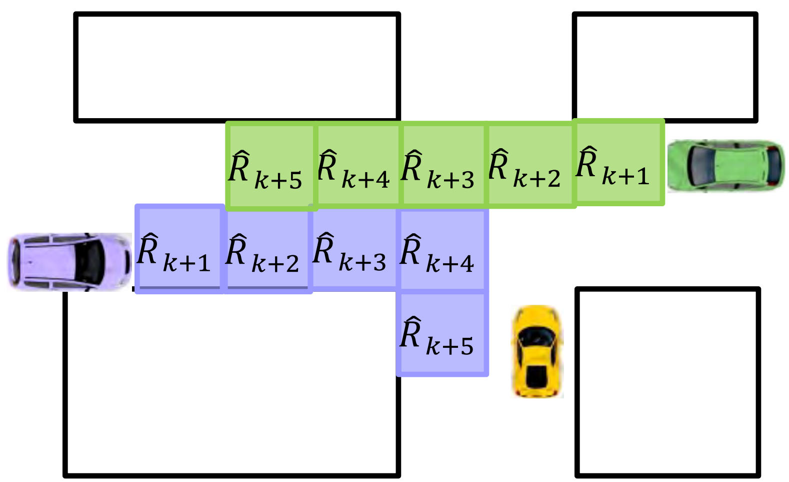
$$\hat{R}_{i,k+d} = [\hat{r}_{s_{k+d}}]$$

$$\mathbb{R}_{i,k:k+D} = \{R_{i,k}, \hat{R}_{i,k+1}, \ldots, \hat{R}_{k+D}\}$$

The agent makes decision based on $\mathbb{R}_{i,k:k+D}$

Expected Rewards

- The expected reward for each position



- A yellow car calculate the expected reward with each position (state s_k) with the probability p_a

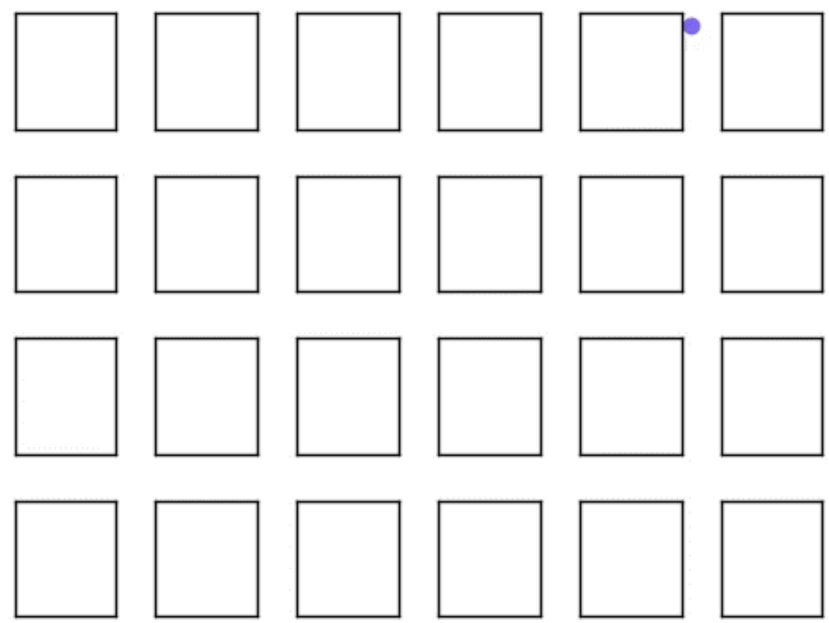
$$\hat{R}_{i,k+d} = [\hat{r}_{s_{k+d}}]$$

$$r_{s'} = \mathbb{E}[r_{s'} | s_{k+d-1} = s'] + \mathbb{E}[r_{s'} | s_{k+d-1} = s]$$
$$= p_{stay} r_{s'} + \sum_a p_a r_s$$

$$a \in \mathcal{A} - \{stay\} = \{forward, left, right\}$$

Ideal V2V Communication

- Manhattan Model Street (M=4, N=6, b=5)
- 20 cars coming into the street
- Communication range r=3

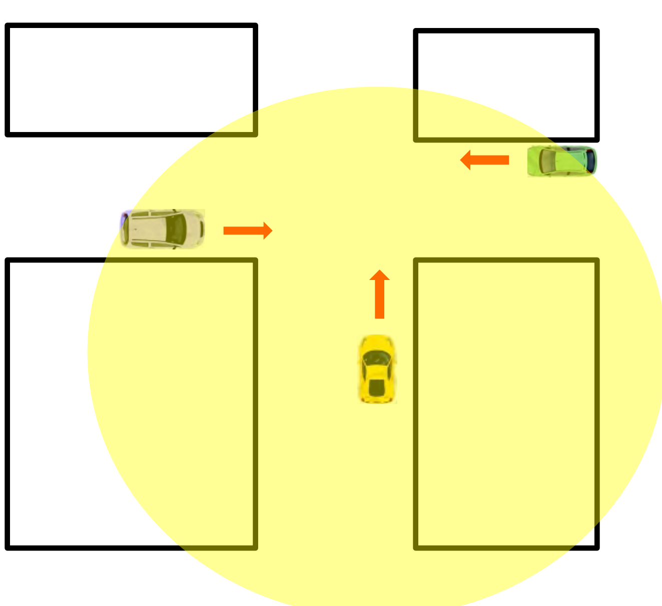


- No Connection
- Connected with other cars

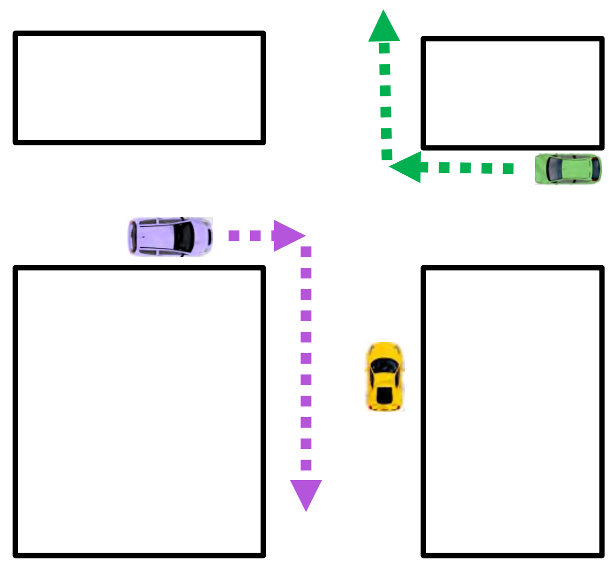
UNIVERSITY OF SOUTH FLORIDA
COLLEGE OF ENGINEERING
Department of Electrical Engineering

Ideal V2V Communication

- Assuming cars can communicate with each other within the communication range r
- In the simple scenario of V2V, cars will have two kind of additional information

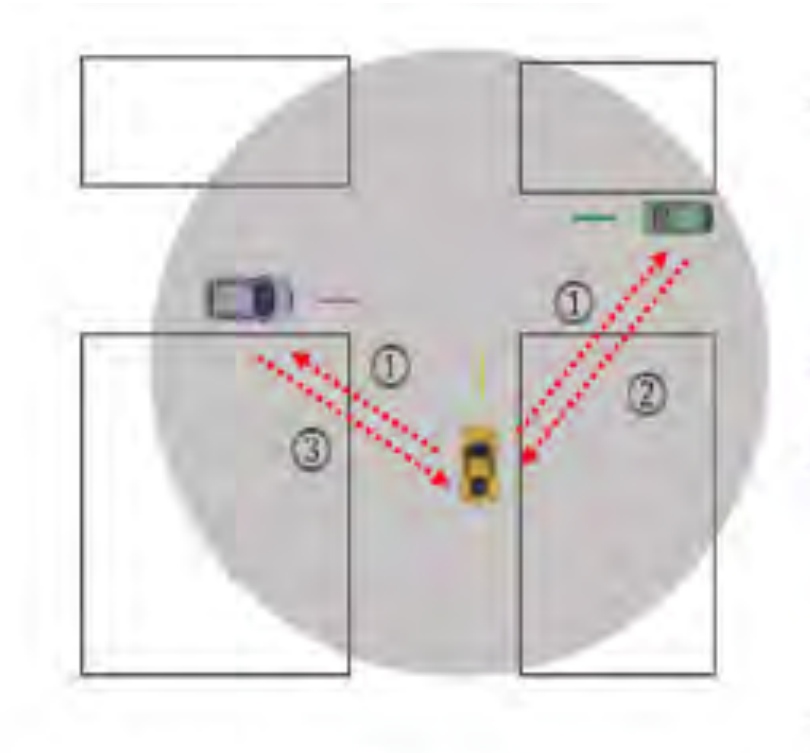


Within the communication range r, the yellow car will recognize the other cars' positions



The yellow car will get the information about other cars' movement (directions)

Ideal V2V Communication



- The i-th AV probes other vehicles' reward maps at time k if any ($j \in \mathbb{I}_k$) AV within the communication range, the i-th successfully receives $\mathbb{R}_{j,k:k+D}$

$$\mathbb{R}_{i,k:k+D} \leftarrow \mathbb{R}_{i,k:k+D} \bigcup_{j \in \mathbb{I}_k} \mathbb{R}_{j,k:k+D}$$

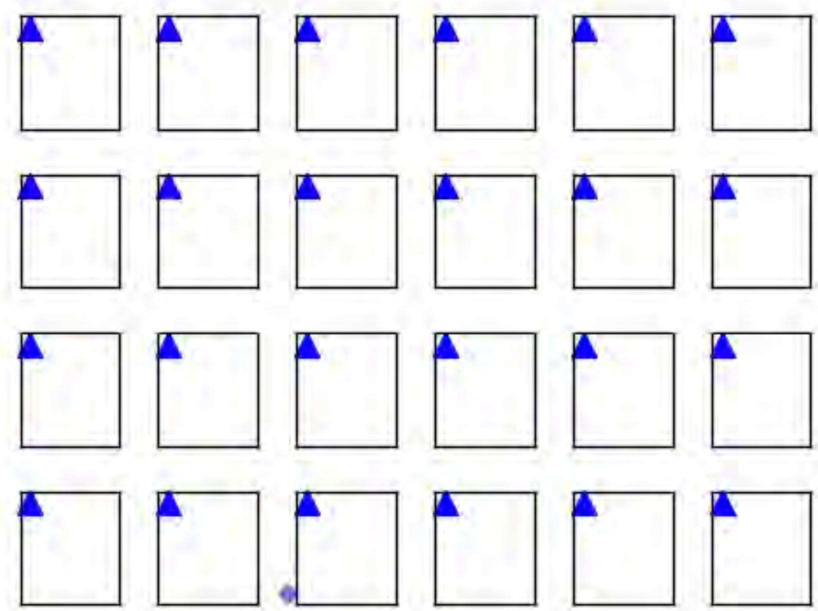
- No matter receiving from other AVs or not, the i –th AV broadcast own reward map $\mathbb{R}_{i,k:k+D}$

$$\mathbb{R}_{j,k:k+D} \leftarrow \mathbb{R}_{j,k:k+D} \bigcup \mathbb{R}_{i,k:k+D}$$

USF UNIVERSITY OF SOUTH FLORIDA
COLLEGE OF ENGINEERING
Department of Electrical Engineering

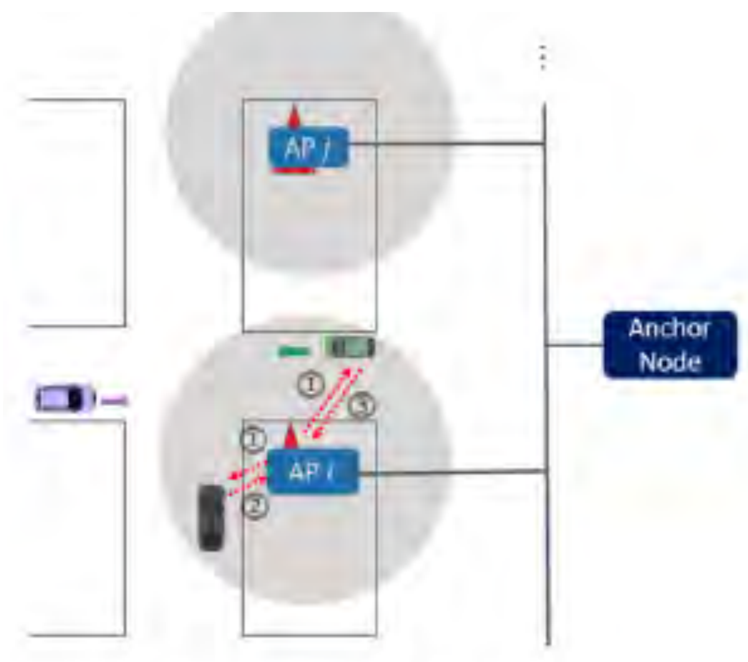
Ideal V2I2V Communication

- Manhattan Model Street (M=4, N=6, b=5)
- 20 cars coming into the street
- Communication range r=3



● No Connection
● Connected with AP
▲ AP: no connection
▲ AP: connecting with cars

Ideal V2I2V Communication



- The network infrastructure (NI) relays the reward maps $\mathbb{R}_{AP_m,k:k+D}$ through APs $m \in M$

$$\mathbb{R}_{AP_{m,k:k+D}} \leftarrow \bigcup_{i \in \mathbb{M}_k} \mathbb{R}_{i,k:k+D}$$

$$\mathbb{R}_{NI_{k:k+D}} \leftarrow \bigcup_{m \in M} \mathbb{R}_{AP_{m,k:k+D}}$$

$$\mathbb{R}_{i,k:k+D} \leftarrow \mathbb{R}_{AP_{m,k:k+D}} \leftarrow \mathbb{R}_{NI_{k:k+D}}$$

What is expected

- Without communication, AV needs to predict the other AV's moving from observation
- For V2V and V2I2V, within the communication range, AV's can swap the accurate information about other vehicles movement

⇩

- Average delays (from the driving time to the destination) can be reduced by V2V and V2I2V communication
- V2I2V communication will improve the performance with smaller range

Multiple Access With Limited Resource

ALOHA with 1 channel

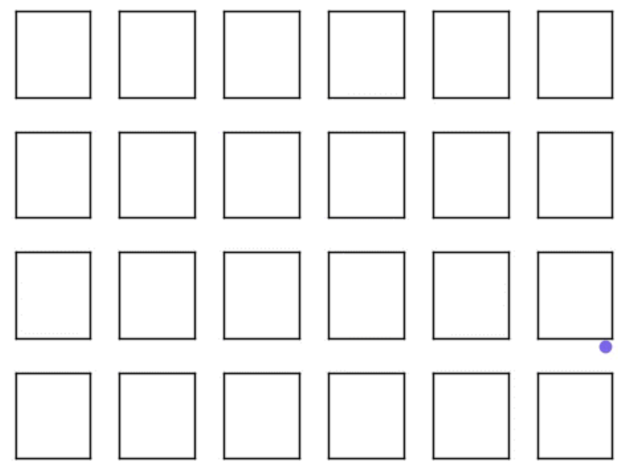
- If there are the packet to transmit, then transmit the packets
- There are collisions when the multi-access occurs

For Vehicle Communication

- If the channel is idle, the car broadcasts own reward map $\mathbb{R}_{\uparrow}$ within communication range *r*
- If the channel is busy, the car waits for receiving other cars reward map $\mathbb{R}_{\downarrow}$
- If the collisions occurs, that means the communication fails (no communication mode)

Multiple Access With Limited Resource

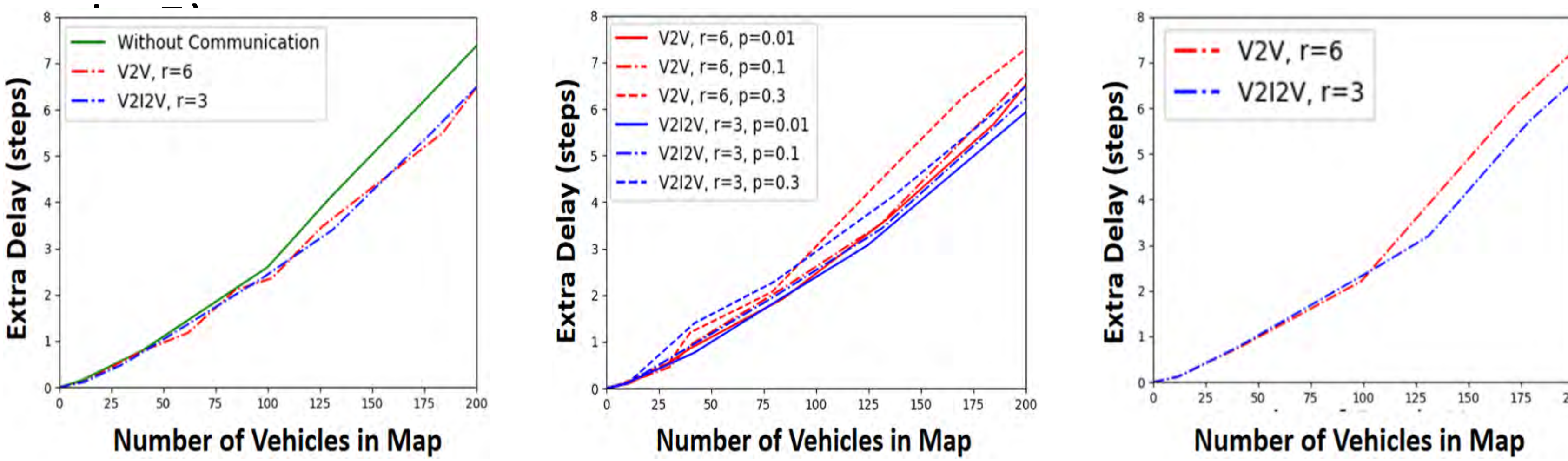
- Manhattan Model Street (M=4, N=6, b=5)
- 20 cars coming into the street
- Communication range r=3



- No Connection
- Connected with other cars
- Collision occurs

Communication Enhances Multi-Agent Systems: New Frontier of AI Computing

- Manhattan Model Street (M=4, N=6,



(a) MAS of RL with Ideal Wireless Communication (b) Message Errors Degrade Performance of MAS (c) Multiple Access Communication by rt-ALOHA

Networked AI

Multiple Access of rt-ALOHA

Modification of slotted ALOHA is required to support RL, named as real-time ALOHA (rt-ALOHA) due to the nature of AI/ML

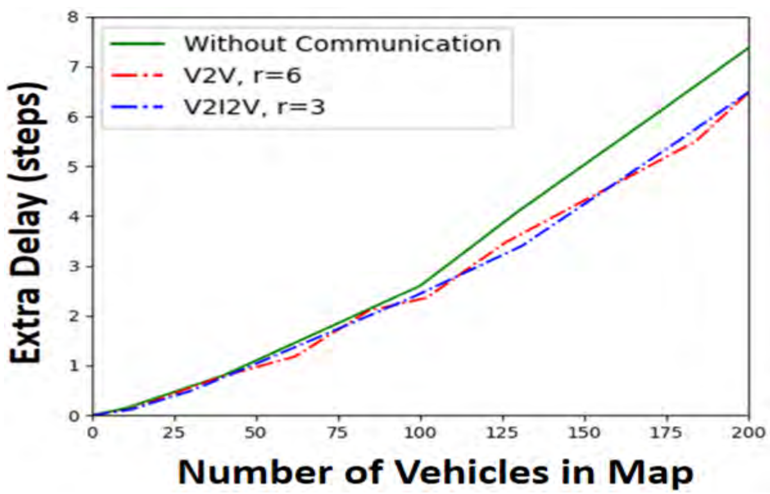
- Data is useless once passing the required networking or communication latency ➔ age of information

- When the channel is busy, the agent (i.e. AV) is ready to receive immediately
- When the channel is idle, the agent broadcasts the message, and ready for receiving from others immediately after transmission without any acknowledgement by the receiving agent(s)
- There is no retransmission and thus backlogging

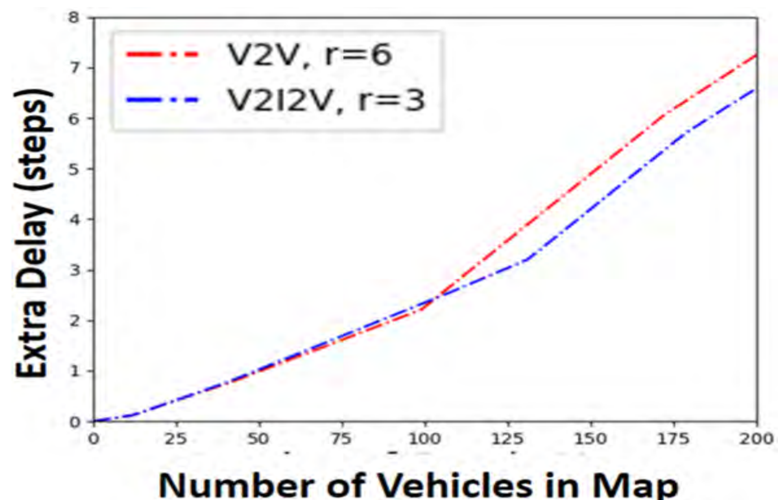
UNIVERSITY OF SOUTH FLORIDA
COLLEGE OF ENGINEERING
Department of Electrical Engineering

Multiple Access of rt-ALOHA

- For V2V and V2I2V communication with one operating frequency channel
- If the channel is occupied by the other agent, goes without communication



Ideal Communication D = 5

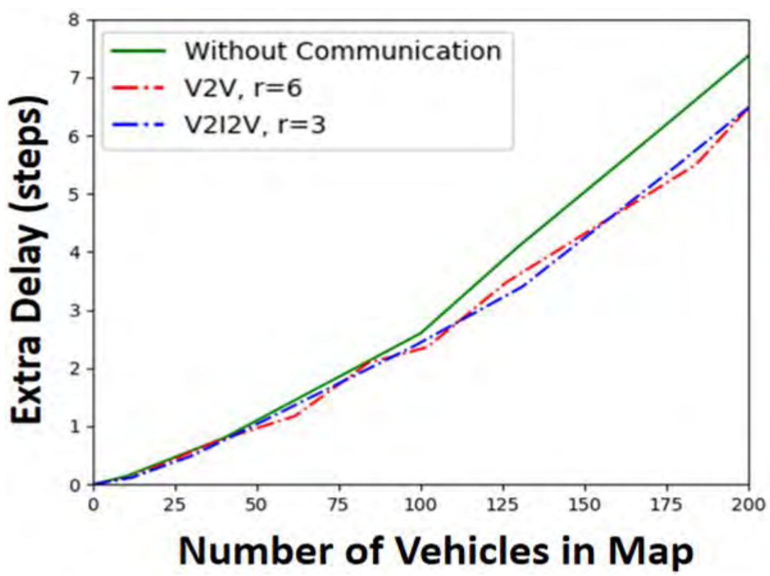


Rt-ALOHA D = 5

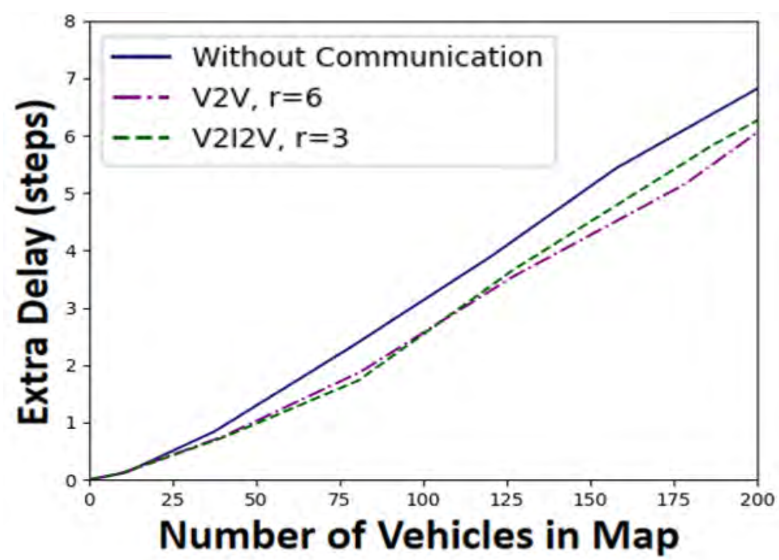
- When V2V and V2I2V communication show the same performance in the ideal communication
- However the longer communication range does not help RL with limited resource unit

Department of Electrical Engineering

Networking can alleviate the load of AI (Learning) computing



D = 5



D = 10

- D means the depth of horizon to obtain the optimal policy in reinforcement learning, while the computational complexity growing exponentially with D
- Collaborative multi-agent systems would be more complicated.

Lessons from Resource-Sharing MAS

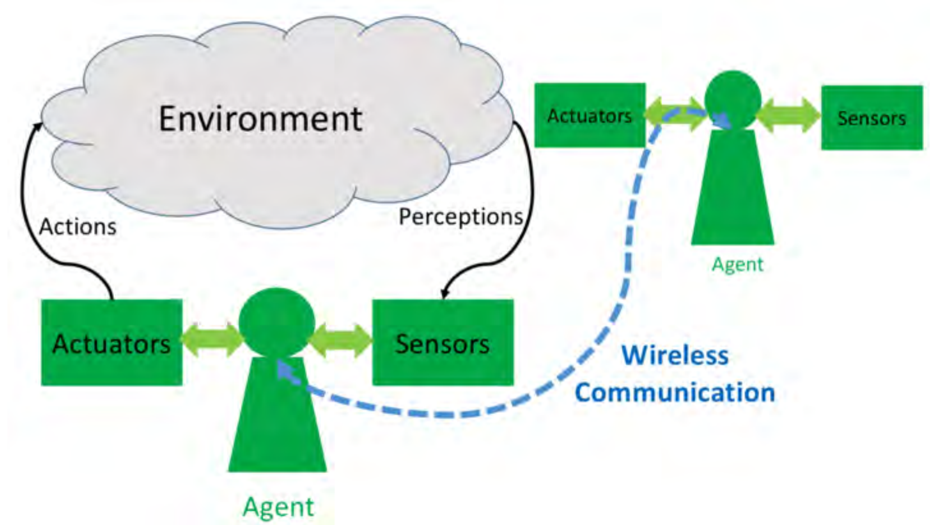
- We observe the how the communication well assists RL, that is, networked artificial intelligence
  - Multiple access consideration suggests small cells.
  - V2I2V communication better assists RL/AI than V2V communication
  - **Appropriate communication alleviates computing load!**
- Real-time (i.e. low-latency) multiple access is more desirable in communication for AI
  - **Age of information** exchanged among agents is a critical factor.
  - **A correctly received message of latency larger than required value in ML is useless in AI.**
  - **Fundamentally different from H2H communication.**

Collaborative multi-agent systems [IEEE ICC 2019]

WIRELESS ROBOTIC COMMUNICATIONS

Networked Multi-Agent Systems

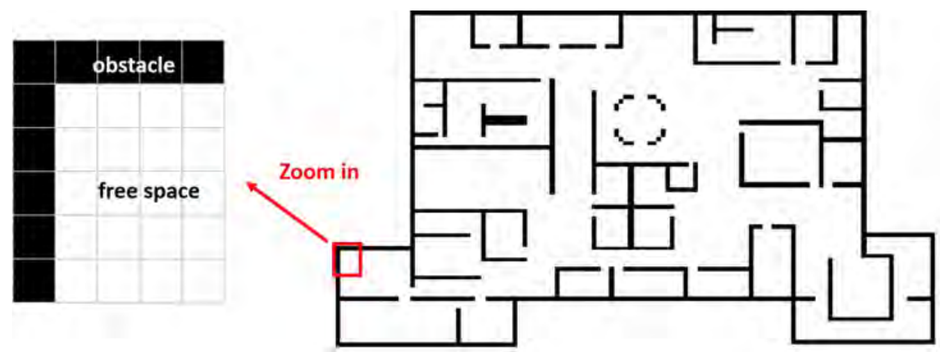
- AI/ML for communications emerges as a new technological frontier.
- However, the impacts of wireless communication on multiple AI agents/robots are rarely known.
  - Multiple AI agents forms a multi-agent system (MAS).
  - MAS of networking is a networked MAS.
  - Eisaku Ko and K.-C. Chen (IEEE GC'18) showed wireless communication enhances resource sharing MAS.



Robots (Agents) of Wireless Communication Capability

Wireless Robotic Communications

- One of the most critical application scenarios is collaborative robots working together toward a common mission in distributed computing
  - Collaborative MAS
  - Smart factory, service robots, etc.
- Goal: To comprehend the role of wireless robotic communication in collaborative MAS
  - Toy example: collaborative robots to clean the floor as right hand layout



Toy Example: Floor plan of the cleaning area, where the area consists of 6760 free space grids and 1227 obstacle grids

System Setup

- Each robot can move up, down, left, and right, in one unit of time. The cleaning task of one tile/grid can be done within the same time unit. These automated cleaning robots share the same mission (i.e. to clean the entire floor) but each of them executes on own intelligence without any centralized controller to manage their actions, as a **collaborative MAS**.
  - No Public Reference: The size and shape of the target area is **unknown** to robots (i.e. agents). In other words, the map of the target area is not available to agents.
  - Localization: Each agent does not know its location at the beginning and must explore to establish its private reference (i.e. own but incomplete map).
  - Each agent equips appropriate sensors and localization algorithm to tell each tile is to-be-cleaned, being cleaned before, and a block. Each agent executes its own learning and decision, which will be modeled as reinforcement learning [9].
  - The time to complete the mission of cleaning the entire (or certain percentage of) floor is used as the performance index of such AI system/MAS.

UNIVERSITY OF SOUTH FLORIDA
COLLEGE OF ENGINEERING
Department of Electrical Engineering

Agent's Private Reference

- Let $g_{p,q}$ label the grid

- Obstacles: A fully occupied grid that is not able to let the robot traverse. Obstacle grids can be represented by $\mathcal{M}_{obs}$.
- Unvisited (uncleaned): A grid that is covered with dirt but have not yet been cleaned. The set of unvisited grids is $\mathcal{M}_X$.
- Visited (cleaned): A grid that has been cleaned and doesn't need to be visited again. The set of visited grids is $\mathcal{M}_O$. (Note that unvisited grids and visited grids are both *free space*, which is denoted by $\mathcal{M}_{free}$.)

- The state of agent u_i is denoted by its location. The state of next instance

$$y^i_{t+1} = \begin{cases} g^i_{j,k} & \text{if } A_t = stay \\ g^i_{j,k+1} & \text{if } A_t = forward \\ g^i_{j,k-1} & \text{if } A_t = back \\ g^i_{j-1,k} & \text{if } A_t = left \\ g^i_{j+1,k} & \text{if } A_t = right. \end{cases}$$

- The reward structure

$$R_{t+1} = \begin{cases} R^+ & \text{if } g_{a,b} \text{ has not been cleaned} \\ R^- & \text{otherwise.} \end{cases}$$

Agent's Reinforcement Learning

- Q-learning due to noisy sensing (with probability p_e) to form the belief

$$P\{\widetilde{R}(g) = R^-|g \text{ dirty}\} = P\{\widetilde{R}(g) = R^+|g \text{ cleaned}\} = p_e$$

$$P\{\widetilde{R}(g) = R^-|g \text{ cleaned}\} = P\{\widetilde{R}(g) = R^+|g \text{ dirty}\} = 1 - p_e \tag{4}$$

- ε-greedy for exploration due to unknown floor map
- TD-n appears to be a good version of RL but actually suffers from lacking of public reference (map).

Single Agent's RL Algorithm

1) Agent is randomly deployed at a free space grid which is defined as coordinate $(0,0)$ in its private reference $\mathcal{M}_t^i, t=0$.
2) Perceives 4 surrounding grids. For all a in action set $\mathcal{A}(b_t^i)$, initialize $Q(b_t^i, a) = Q_0$ if $Q(b_t^i, a)$ has not been defined. Q_0 is just an initial value that could be set at any value.
3) Calculate action-value function using $\widetilde{R}_{t+1}$ in reward map. The tilde over $\widetilde{Q}(b,a)$ indicates that it is not the real action-value but the estimated one. $\forall a \in \mathcal{A}(b_t^i)$,

$$\widetilde{Q}(b_t^i, a) \leftarrow Q(b_t^i, a) + \alpha[\widetilde{R}_{t+1} + \arg\max_{a'} Q(b_{t+1}^i, a') - Q(b_t^i, a)]$$

4) Let the optimal action be $a^* = \arg\max_a Q(b_t^i, a)$. Choose action A_t following ϵ-greedy policy, that is

$$A_t = \begin{cases} a^* & \text{with probability } 1-\epsilon \\ a \neq a^* & \text{with probability } \frac{\epsilon}{|\mathcal{A}(b_t^i)|-1}. \end{cases}$$

5) Operate action A_t, transits to state b_{t+1}^i, and receive reward R_{t+1}. Update the action-value function

$$Q(b_t^i, a) \leftarrow Q(b_t^i, a) + \alpha[R_{t+1} + \arg\max_{a'} Q(b_{t+1}^i, a') - Q(b_t^i, a)]$$

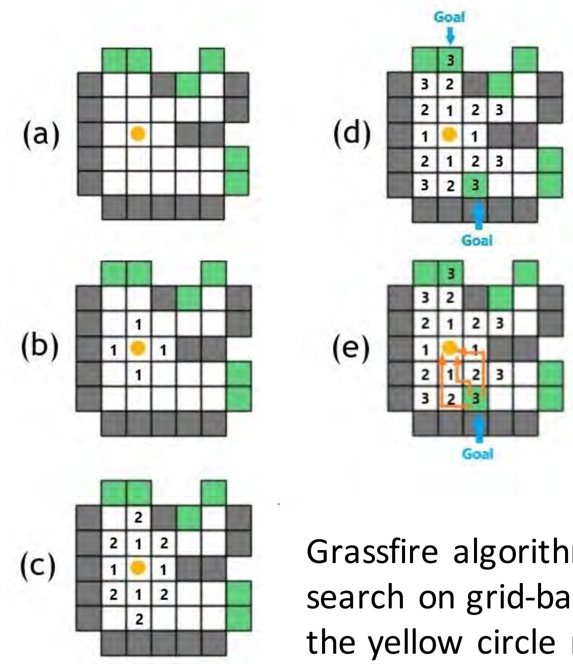
6) $t \leftarrow t+1$. If all rewards on reward map is R^-, meaning that all grids have been cleaned, terminate. Otherwise, go back to step 2)

Robot Panning

- Purely relying on reinforcement learning is very ineffective to complete a large-scale mission without public reference, which suggests the importance of planning algorithm
  - Fixed depth planning: agent behavior and retrieving limited information in reward map. That is, the agent first examines all grids within Manhattan distance *d* and goes toward the nearest grid that has positive reward.
  - Conditional exhaustive planning: switching conditions for agent to start exhaustive planning on reward map, to strike balance in exploitation and exploration when public reference is not available.
    - More importantly, exploration helps to establish a private reference and private reward map which become valuable resource that can be provided for others and for its future decision.

Conditional Exhaustive Planning

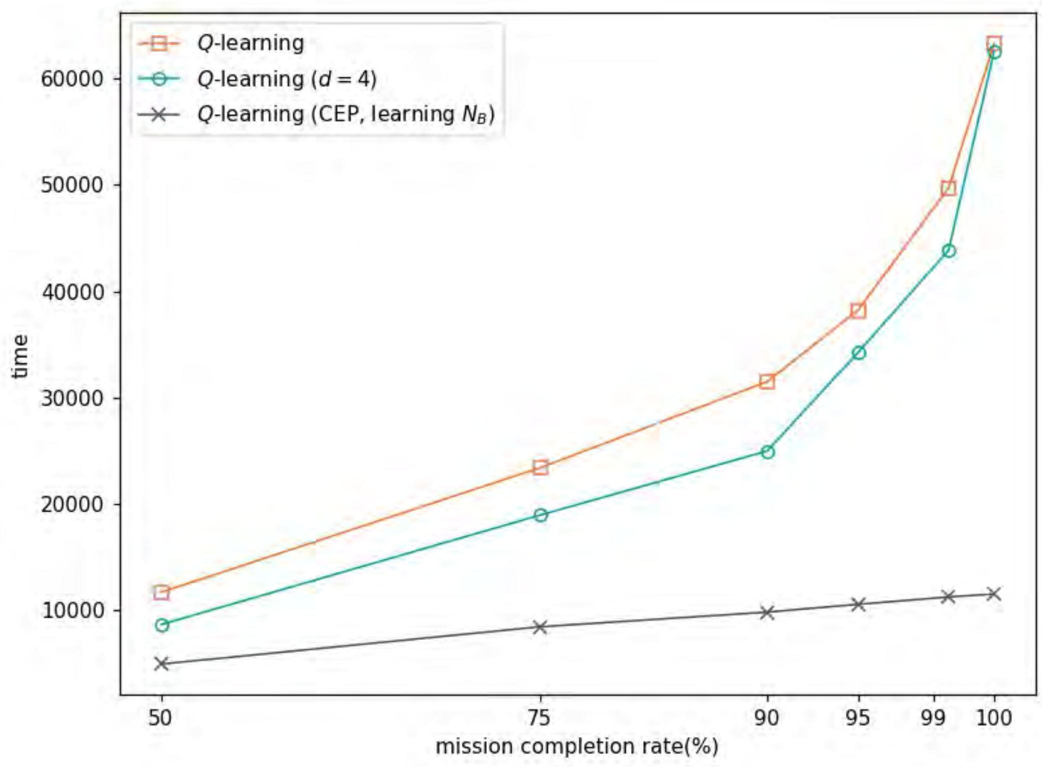
- Conditions to adopt planning: Since exploration and planning are both likely to benefit the agent's task, we consider a block size of N_B.
- Grassfire algorithm: a breadth-search first methods on grid-based graph, can effectively search and construct a path to the goal on a graph
- Learning N_B



Grassfire algorithm: breadth-first search on grid-based map, where the yellow circle represents the current position from which we try to find a nearest goal, i.e. green grids.

UNIVERSITY OF SOUTH FLORIDA
COLLEGE OF ENGINEERING
Department of Electrical Engineering

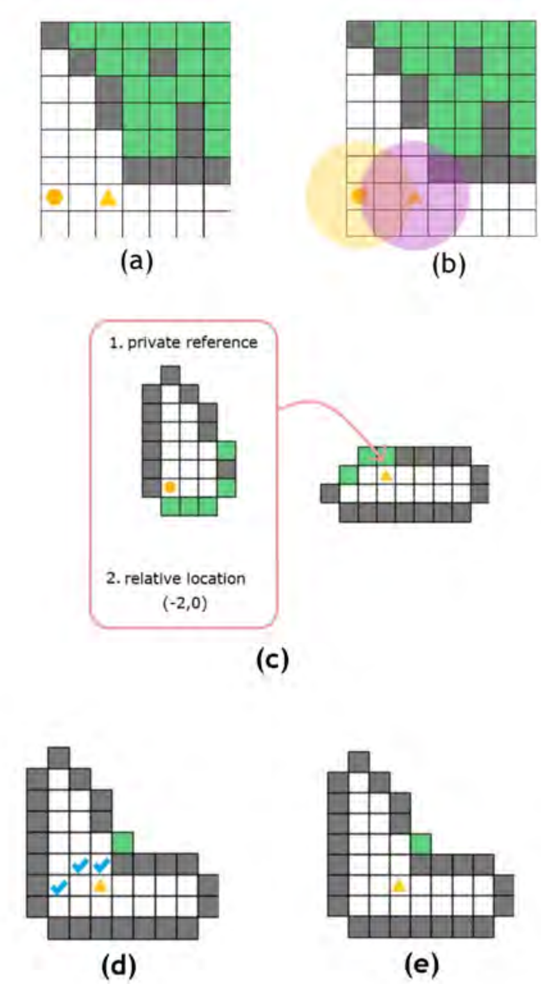
Q-learning incorporating with planning and localization



UNIVERSITY OF SOUTH FLORIDA
COLLEGE OF ENGINEERING
Department of Electrical Engineering

Information Exchange and Integration for Collaborative Agents

- each agent shall transmit its relative location and private reference $M_{i,t}$ (i.e. map being explored up to now), with information of grids being visited and sensed (i.e. corresponding state values).
- Furthermore, as indicated in earlier research, state-value function with reward map shall be also sent.
- At the receiver end(s), after obtaining external private reference ($M_{j,t}$, $j \neq i$) and experience, the robot will update the original private reference M_i



Multiple Access: *p*-persistent rt-ALOHA

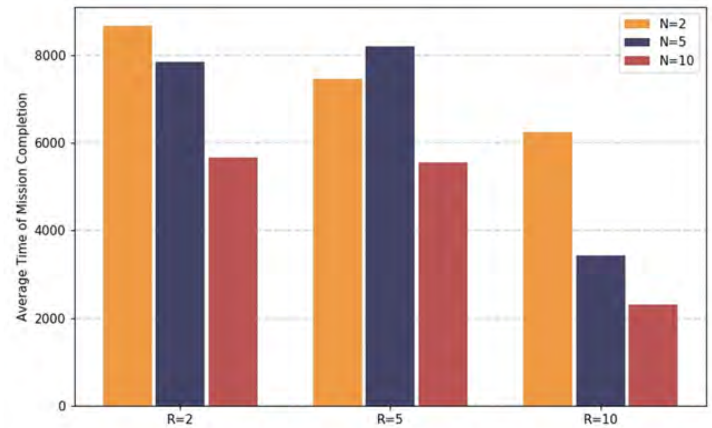
Latency is so critical for information exchange among collaborative agents to suggest real-time ALOHA

- When the channel is busy, the agent (i.e. cleaning robot) is ready to receive immediately.
- When the channel is idle, the agent broadcasts the message of desirable content, then immediately turns ready to receive from others right after transmission without any acknowledgement.
- There is no retransmission and thus backlogging.

Borrowing the concept from CSMA

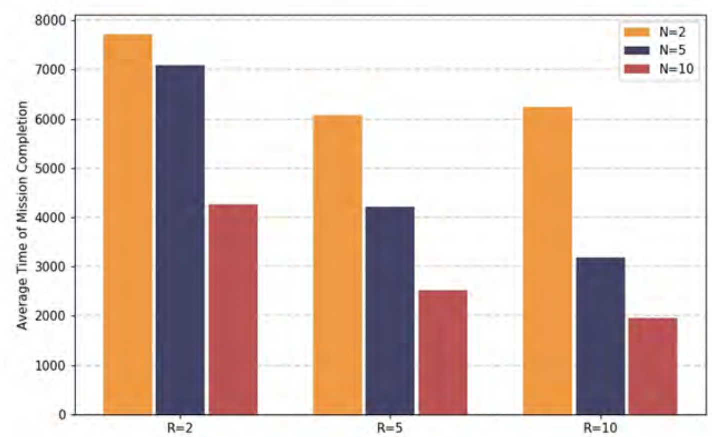
- Proactive: if the agent senses other agents are within its communication range and the channel is not busy, agent broadcasts messages with probability p_p.
- Reactive: When the multi-access channel is busy, the agent stays at the reactive mode, ready to receive other's broadcast. When agent senses other agents are within its communication range, agent stays at the reactive mode with probability $1 - p_p$.

Collective Performance of Collaborative Agents Using *p*-persistent rt-ALOHA



(a)

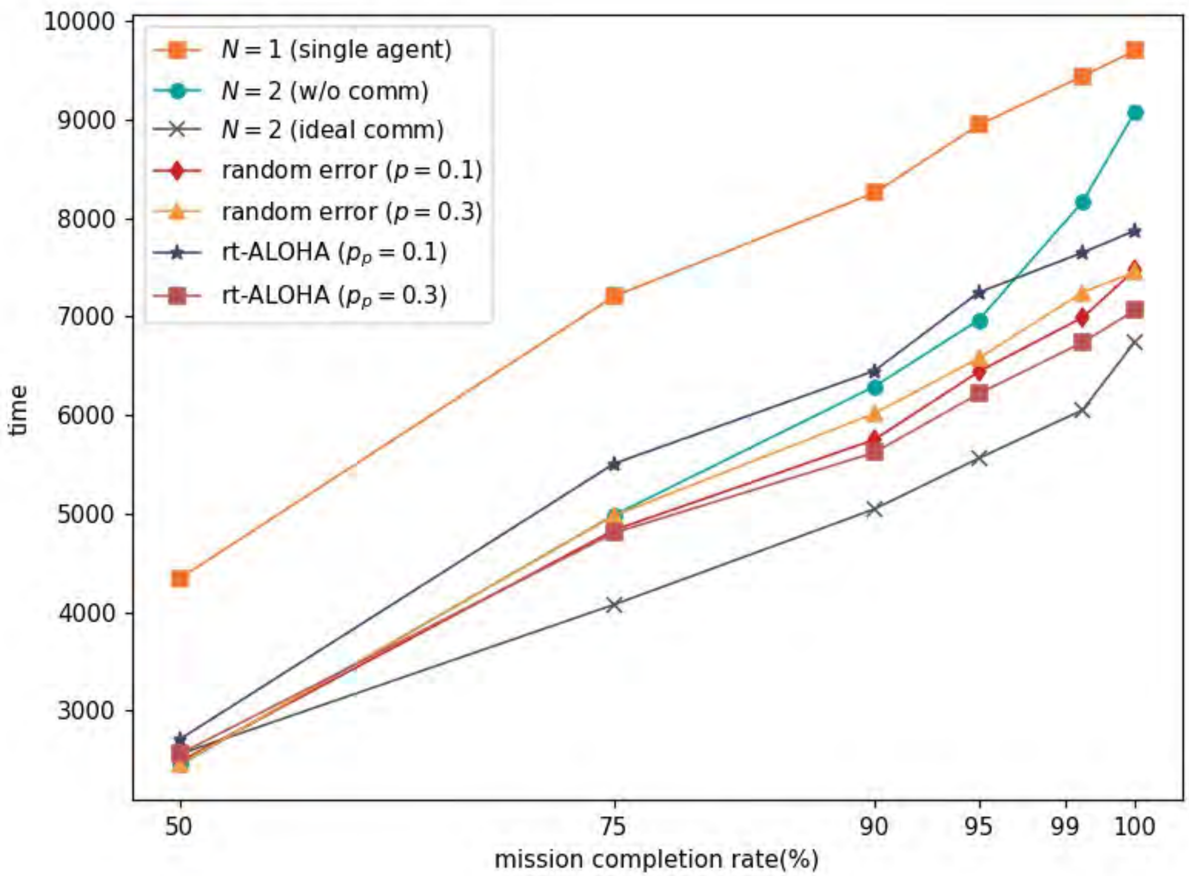
(a) with $p_p = 0.1$
(b) with $p_p = 0.3$ for
N =2 (yellow), 5 (blue), 10 (red).



(b)

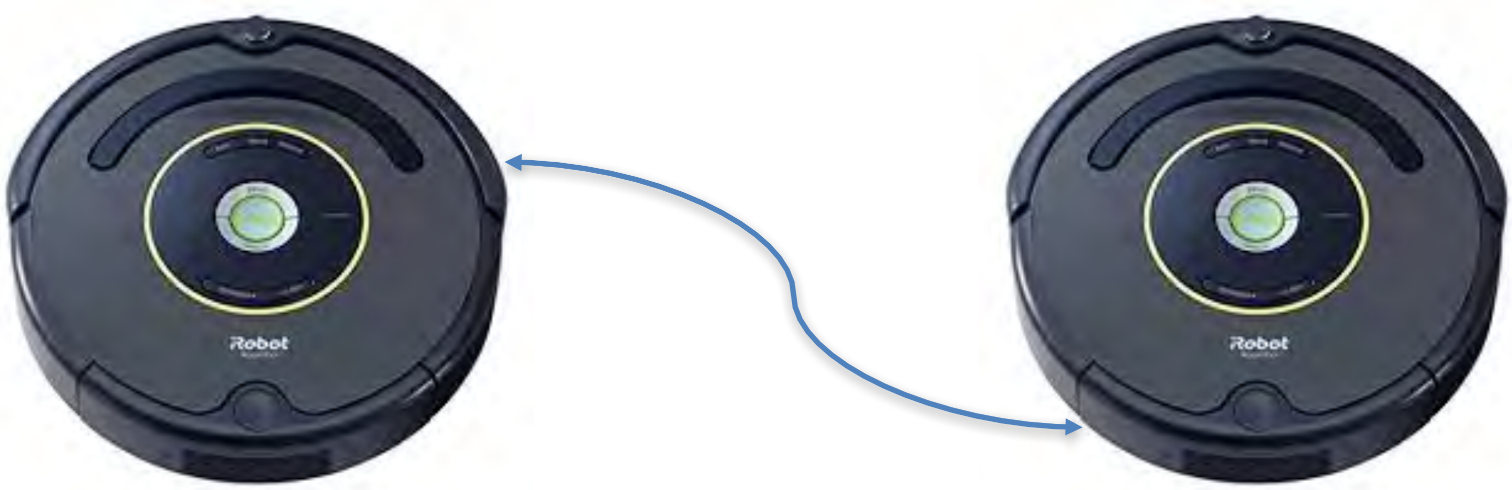
UNIVERSITY OF SOUTH FLORIDA
COLLEGE OF ENGINEERING
Department of Electrical Engineering

Wireless Communications Greatly Enhances Collective Performance of Collaborative Agents (MAS)



Department of Electrical Engineering

The second cleaning robot is useless if they can not communicate each other!



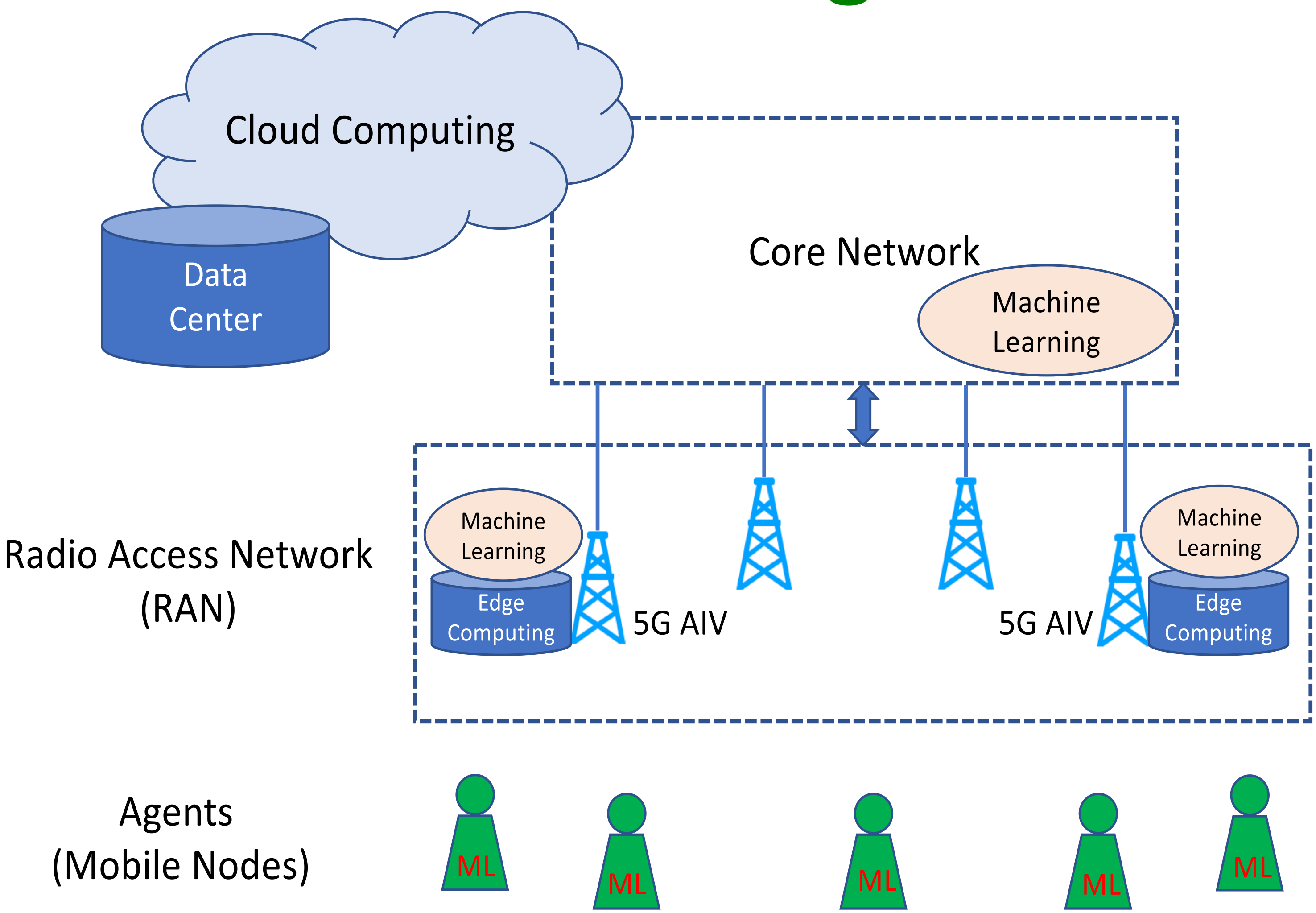
Lessons

- Wireless communications and networking can significantly enhance the collective performance of collaborative robots/agents/MAS (i.e. AI, while AI computing will be limited by hardware)
  - Latency can greatly degrade the useful information among agents
  - rt-ALOHA is therefore useful in ultra-low latency communication and system reliability
  - *p*-persistent provides a means to adjust multiple access communication among agents/robots
  - Application scenarios include smart factory/manufacturing, service robots, and autonomous vehicles.
- Wireless robotic communications, a new kind of machine-to-machine communications, emerges!
  - Again, latency and consequently age of information/data is much more important than throughput in networking AI.

ML/AI and WC

- ML/AI enhances WC
- ML/AI enables WC
- WC enables ML/AI
  - E. Ko, K.-C. Chen, "Wireless Communications Meets Artificial Intelligence: An Illustration by Autonomous Vehicles on Manhattan Streets", *IEEE Globecom*, 2018.
  - K.-C. Chen, H.-M. Hung, "Wireless Robotic Communication for Collaborative Multi-Agent Systems", *IEEE International Conference on Communications*, 2019.
- Future Network Architecture for AI/ML

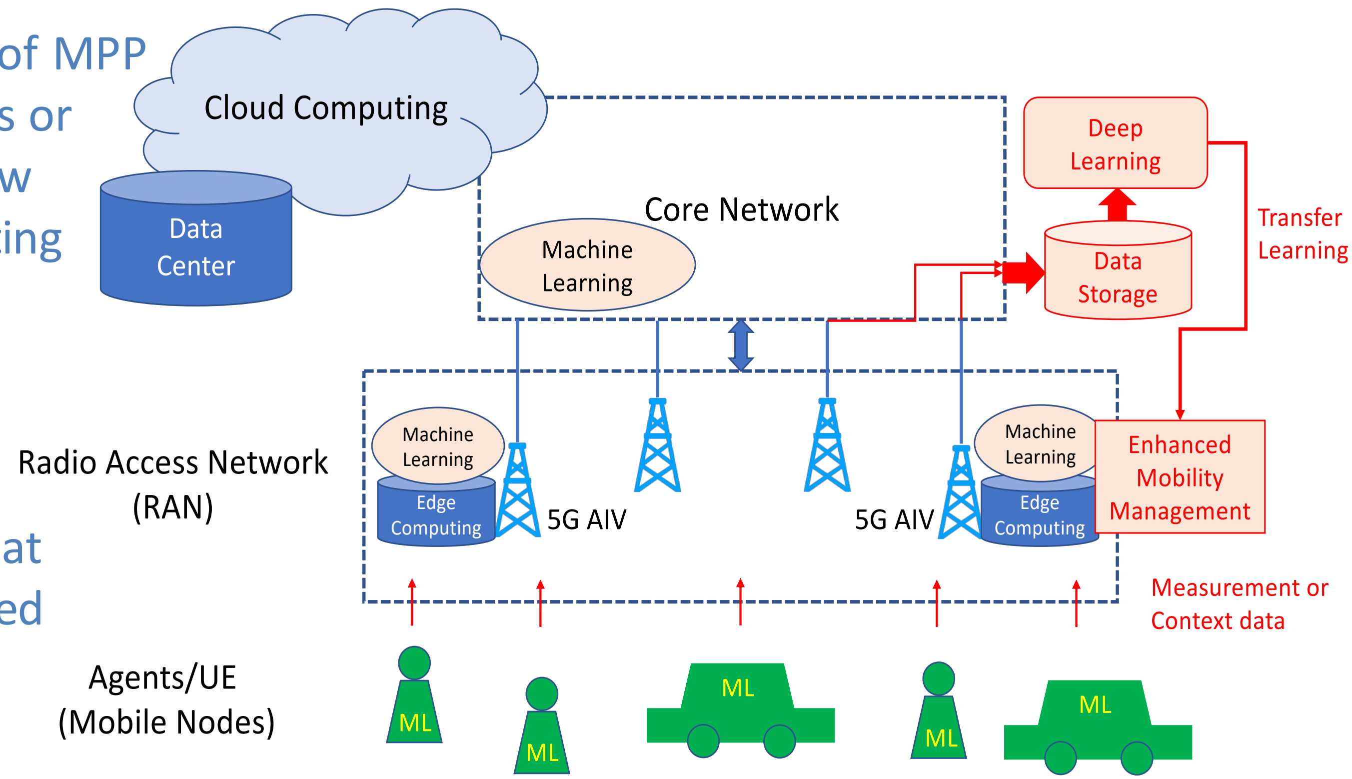
Holistic View of Computing and Networking



Network Architecture of Offline Machine Learning

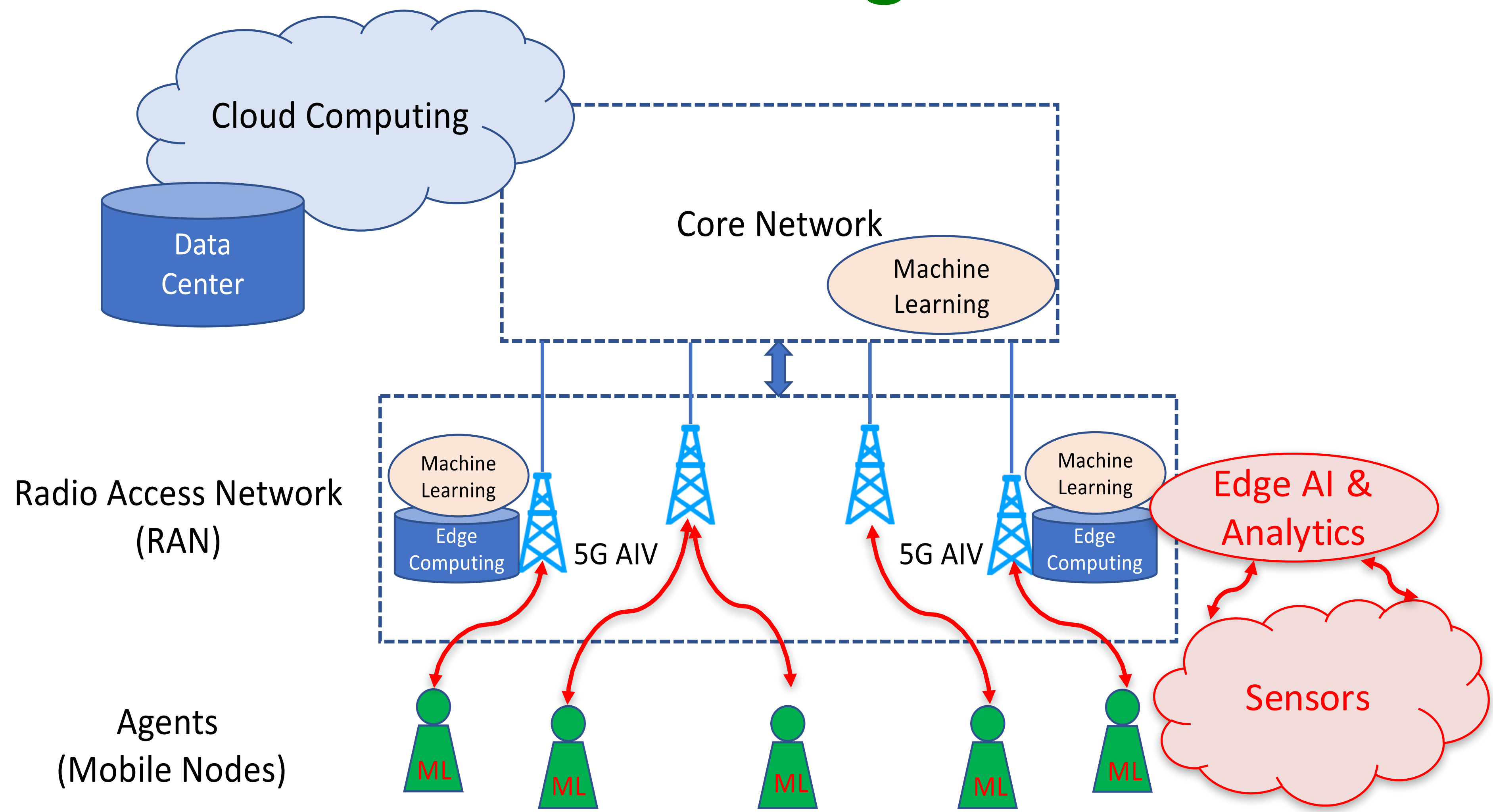
Possible Realization of MPP with New Data Flows or Connections and New Networking/Computing Entities in Network Architecture (in red lines/boxes)

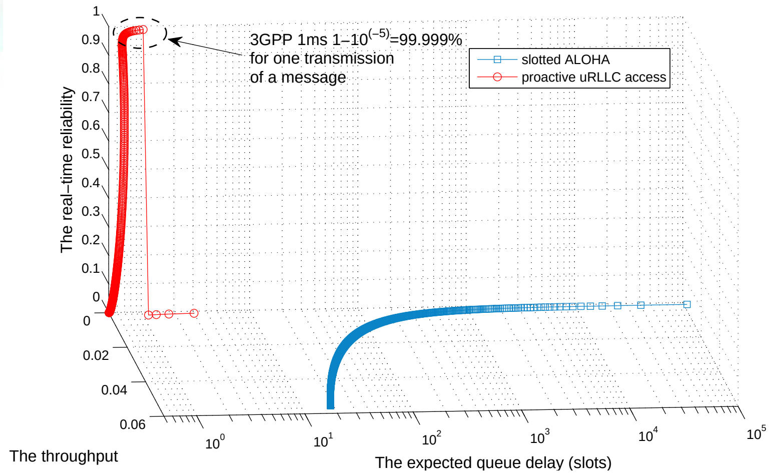
Offline ML means that ML is not directly used for online network functionalities.



UNIVERSITY OF SOUTH FLORIDA
COLLEGE OF ENGINEERING
Department of Electrical Engineering

Holistic View of AI Computing and Networking





Not just autonomous vehicles,

but also smart factory, service robotics, …

COMMUNICATION FOR AI:
A NEW TECHNOLOGY PARADIGM