xS3D summer lab seminar

# Multi-agent MDP model for task allocation in persistent robotic security service

Minjun Kim

2017/09/08

# Presentation overview

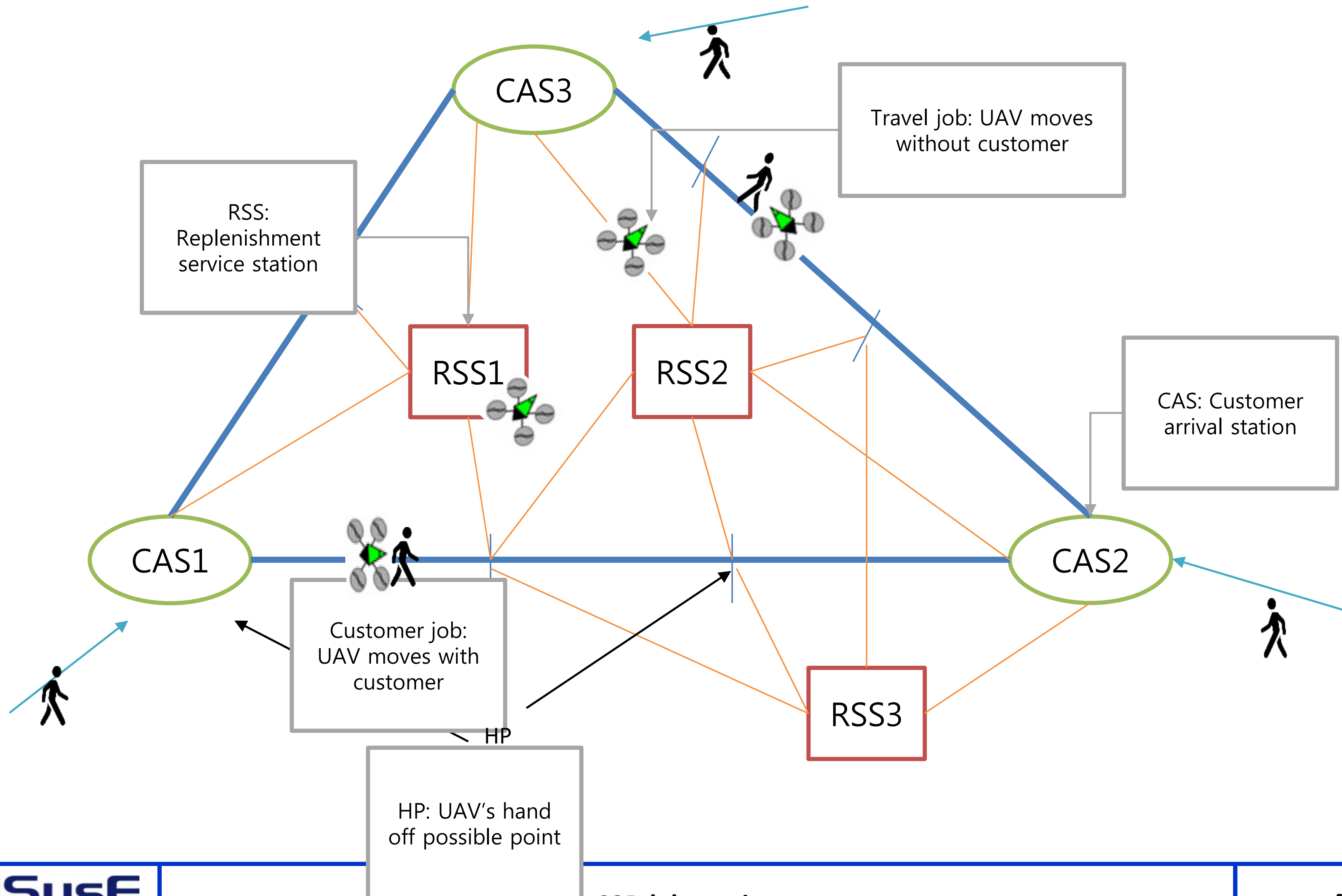
- Introduction
- Markovian modeling of UAV system
- Solution approach
- Numerical analysis
- Concluding remark

# Introduction

- Increasing demand of Unmanned Aerial Vehicles (UAVs)
  - The global market size in 2020 is estimated at 15 billion dollars!!
- Lots of applications of UAV
- Application of this study
  - UAV provides security service to customer

- Develop task allocation method for security system using MDP

# UAV system description



# Contribution

- Markovian modeling of UAV system
  - Generic modeling in stochastic environment
  - Fuel replenishment, hand off, multiple locations, customer distribution
- Develop solution approach of MDP model
  - Dynamic programing approach (value iteration)
  - Greedy heuristic
  - Simplified MDP based heuristic
- Numerical analysis of multi-UAV system
  - Solving and simulation
  - Ensuring optimal solution's validity

# Difficulty of the problem

- Curse of modeling
  - In order to generate stochastic model in real world problem, many assumptions are required it

- Curse of dimensionality
  - If number of UAVs are 7 and maximum number of customers in the system is 6:
    - Number of state: $(58)^7 \times \frac{12!}{6! \times 6!} \approx 2.04 \times 10^{15}$ states
    - There is limit in dynamic programming approach
    - Simulation, approximation is needed

# Generic state of UAV system

- Index for UAV: $i$ , Number of UAVs: $n$
- Index for customer job: $j$, Number of customer jobs: $m$
- Index for jobs and RSS and HP: $k$, Number of jobs and RSS and HP: $p$
- Location of $UAV_i$ : $l_i \in \{Customer\ job\ or\ Travel\ job\ or\ RSS\ or\ HP\ k\}$
- Fuel level of $UAV_i$: $\boldsymbol{f_i \in \{0, \Delta f, 2\Delta f, 3\Delta f, \dots, F_{max} - \Delta f, F_{max}\}}$
- Total number of customers in the system:

$$u_{ij} \in \{0,1\},\ \ w_j \in \{0,1,..,c\},\ \ \sum_{i=1}^{n}\sum_{j=1}^{m} u_{ij} + \sum_{j=1}^{m} w_j \le h$$

- Representation of general state

$$S=\{(l_1, f_1), (l_2, f_2), \dots, (l_n, f_n), w_1, w_2, \dots, w_m\}$$

UAV 1's location

UAV 2's fuel level

Number of waiting customer for customer job 2

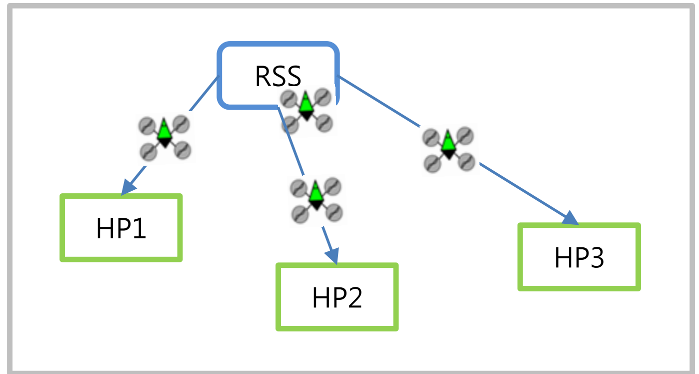
# Problem description of UAV system

- Objective
  - Find the policy (best decision) in each state which maximise the expected value of long term discounted customer job serve rate
    - $V(x) = max_a(\sum_i R_j * n_{ij}(x) + \gamma(\sum_{x'} P_a(x, x')\, V(x'))$
    - $\pi_x = argmax_a\, V(x)$

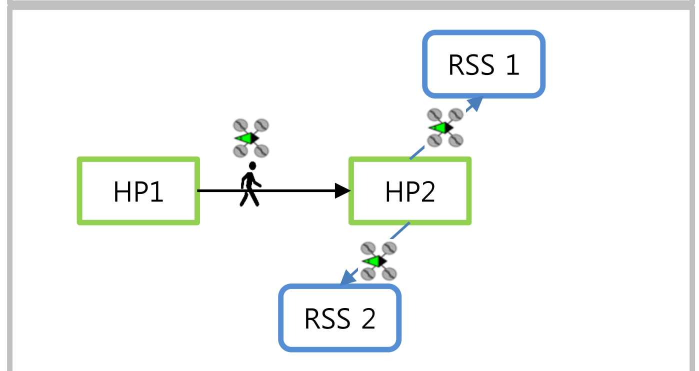
- Constraints
  - UAV's fuel limitation
    - $f_i \in \{0, \Delta f, 2\Delta f, 3\Delta f, \ldots, F_{max} - \Delta f, F_{max}\}$
  - Maximum number of customers in the system
    - $u_{ij} \in \{0,1\},\ w_j \in \{0,1,..,c\},\ \sum_{i=1}^{n}\sum_{j=1}^{m} u_{ij} + \sum_{j=1}^{m} w_j \leq h$

# Decisions

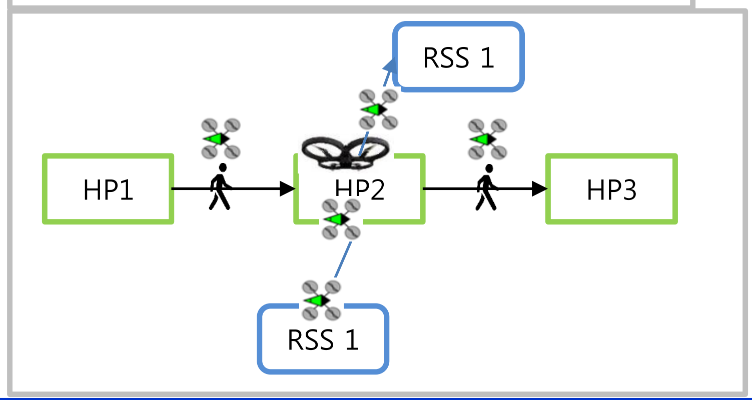
- After fuel replenishment in RSS, UAV moves to which HP?




- After serving customer, UAV moves to which RSS?




- After serving customer, UAV continues to serve that customer or hand off with another UAV  and moves to RSS or stay in HP?



# Greedy heuristic (GDH)

Algorithm: Greedy heuristic

Input: Current state

Output: Each UAV’s decision in current state

1. Classify state (hand off, travel job which is same path as customer job)

2. From all UAVs, If number of possible action is 1, follow that action

3. When deciding the each UAV’s action, concern their current location, fuel level and significant customer from each UAV.  Do not concern other UAV’s information

4. Each UAV moves to provide service to significant customer

# Simplified MDP based heuristic (MDH)

Algorithm: Simplified MDP based heuristic

Input: Current state

Output: Each UAV's decision in current state

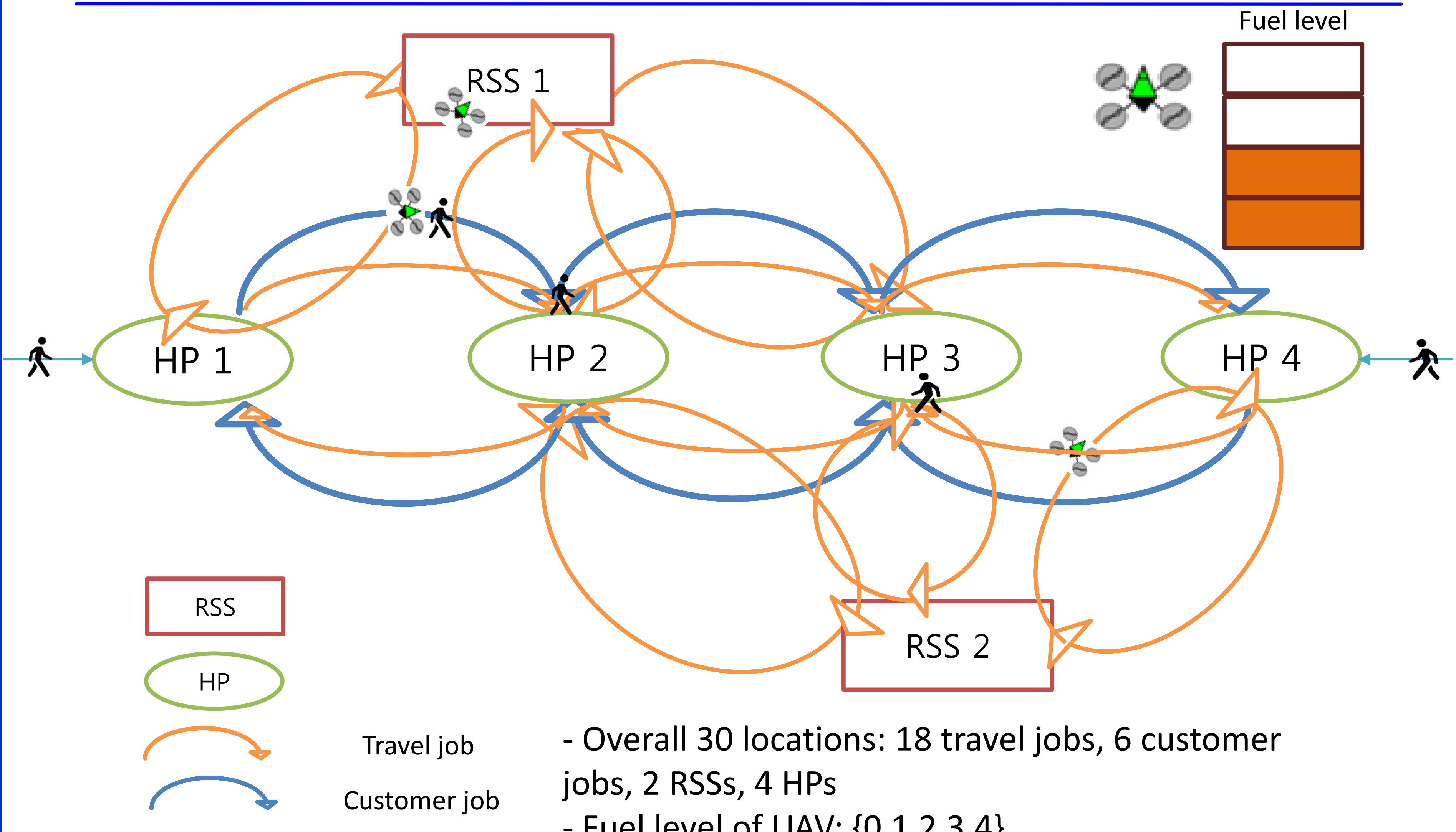
1. Classify state (hand off, travel job which is same path as customer job)

2. From all UAVs, If number of possible action is 1, follow that action

3. Generate simplified MDP models. It consists of main UAV, UAV closest to the main UAV, same customer distribution as the original model

4. Use simplified model`s VI solution to decide each UAV's action

# Simulation system description




- Overall 30 locations: 18 travel jobs, 6 customer jobs, 2 RSSs, 4 HPs
- Fuel level of UAV: {0,1,2,3,4}

# Simulation

- Input variable

| Variable | Customer arrival rate | Customer job serve rate | Travel job serve rate | Fuel replenishment rate for 1 unit | Reward for serving customer |
|---|---|---|---|---|---|
| Symbol | $\boldsymbol{\lambda}_a$ | $\mu_C$ | $\mu_T$ | $\mu_F$ | $r_j$ |
| Value | 0.1 | 0.4 | 0.8 | 0.5 | 1 |
| Distribution | Exponential | Exponential | Exponential | Exponential | Constant |

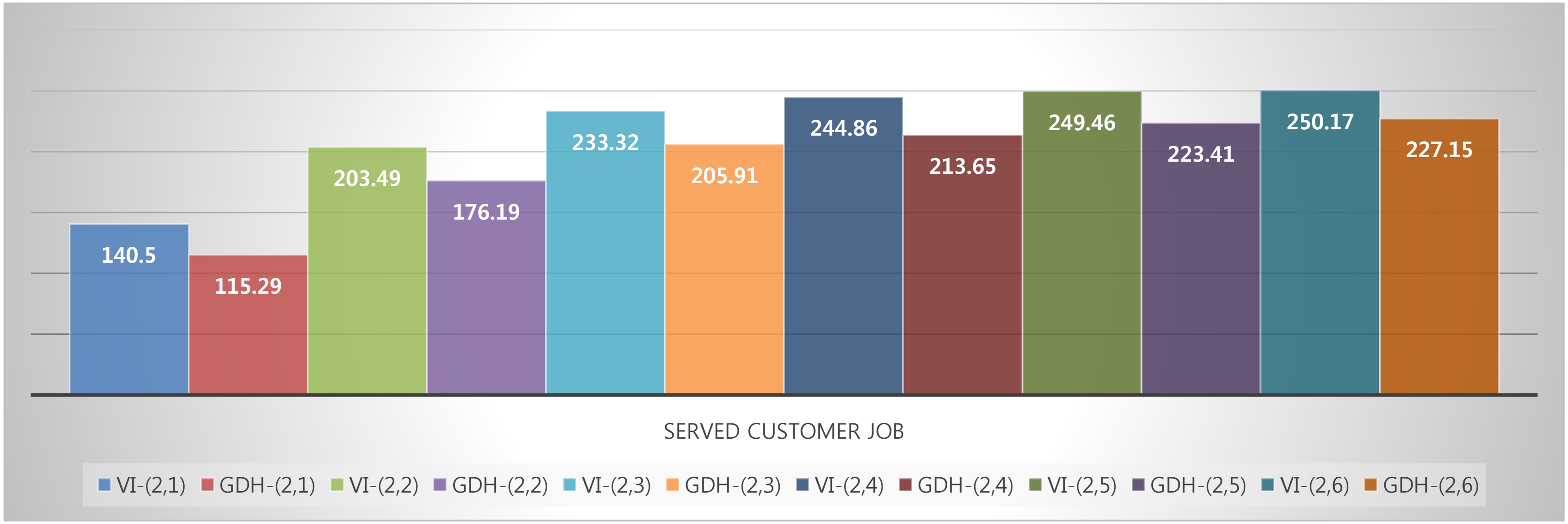
  - Each state's decision is solution of MDP
  - Simulation time limit: 1000
  - Number of simulations: 100
- Output variable
  - Each job's start time and finish time with provided UAV
  - Use the result to get the total number of served customer job within the simulation time

# Logic of simulation

- Set initial state
- Customer is arrived and transition is occurred and UAV provide security service.
- Repeat
  - 1. Set the decision of current state (Based on algorithm{MDH, GDH, VI})
  - 2. From current state, find the feasible transition and find the time the event can occur (Job completion time of each UAV, customer arrival time )
  - 3. Among the feasible event time, the most recent event time`s event is occurred. This time is recorded
  - 4. Transition is occurred and UAV did job, each job`s start time and final time with provided UAV is recorded
  - 5. If event occurred time is bigger than simulation time limit, stop it

# Simulation results: 2 UAVs cases

VI: Value iteration GDH: greedy heuristic

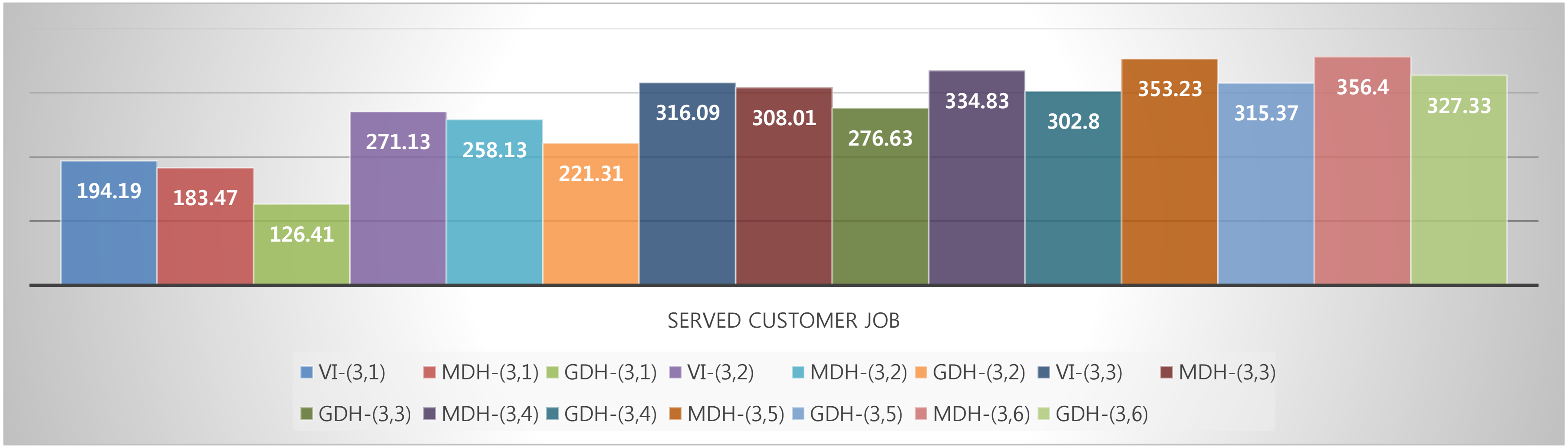



- Optimality gap compared with value iteration (%)

| Model | (2,1) | (2,2) | (2,3) | (2,4) | (2,5) | (2,6) |
|---|---|---|---|---|---|---|
| GDH | 17.94 | 13.42 | 11.75 | 12.75 | 10.44 | 9.20 |

# Simulation results: 3 UAVs cases

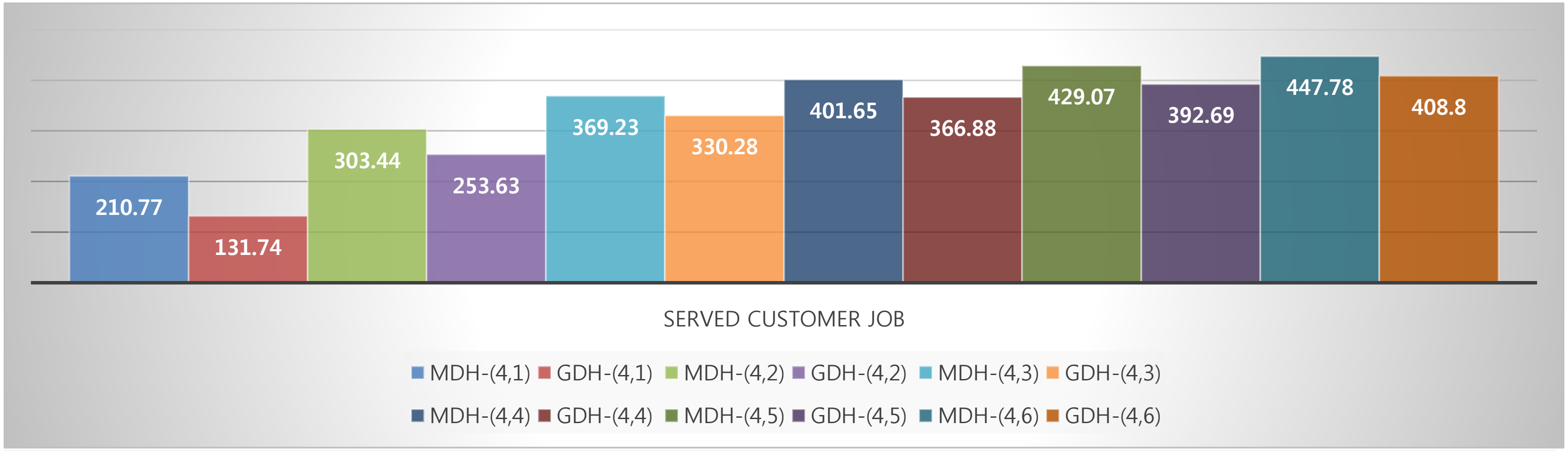
VI: Value iteration GDH: greedy heuristic MDH: simplified MDP based heuristic

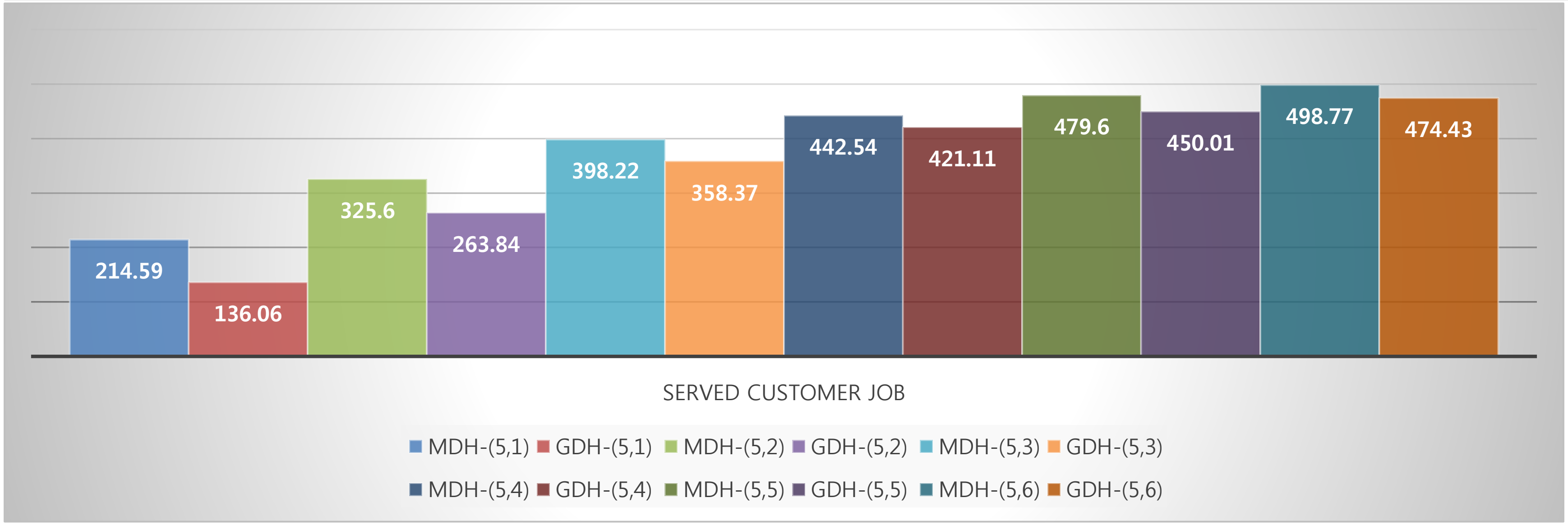



- Optimality gap compared with value iteration (%)

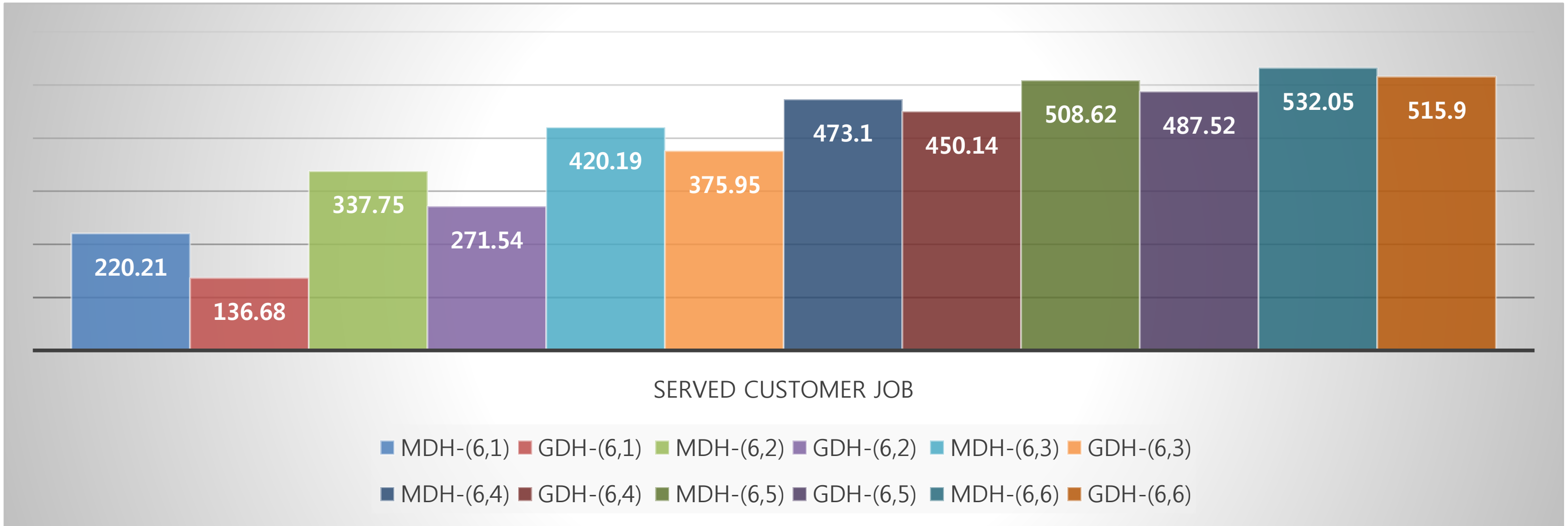
| Model | (3,1) | (3,2) | (3,3) |
|---|---|---|---|
| GDH | 34.90 | 18.37 | 12.48 |
| MDH | 5.52 | 4.79 | 2.59 |

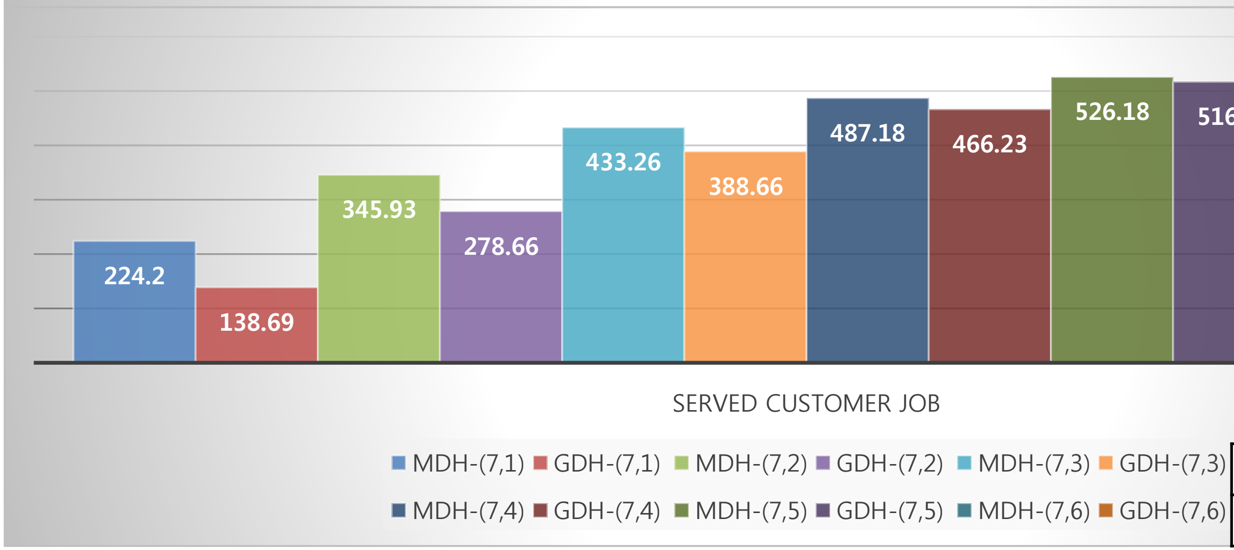
# Simulation results: 4 UAVs, 5 UAVs cases



# Simulation results: 6 UAVs, 7 UAVs cases




$2.04 \times 10^{15}$ states

| | MDH | GDH |
|---|---|---|
| Service rate | 91.2 | 89.8 |

# Summary of result

- There is limitation in dynamic programming approach
  - It can be applied until 3 UAVS and 3 maximum customers case
- Heuristic policy can be appied multi-UAVs case
  - We simulated until 7 UAVs and 6 maximum customers case. In this case MDH, GDH heuristic`s service rate is about 91.2%, 89.8%.
- Simplified MDP based heuristic shows better performance
- As maximum number of customer in the system is increased, GDH heuristic`s performance is increased
- As number of UAVs in the system is increased, the gap between two heuristic is reduced

# Concluding Remark

- Generic MDP modeling of UAV system with random customer arrivals and task durations

- To address computational complexity
  - Efficient representation of states and transitions
  - Greedy heuristic(GDH), simplified MDP based heuristic(MDH)

- Numerical analysis of multi-UAVs system
  - Apply value iteration until 16389408 states, 27 actions case
  - Apply heuristic until $2.04 \times 10^{15}$states, 2187 actions case
  - GDH 15% suboptimal, MDH 4% suboptimal