



VLDB2024
GUANGZHOU



北京航空航天大学
BEIHANG UNIVERSITY

A Data-driven Spatiotemporal Simulator for Reinforcement Learning Methods

Dingyuan Shi¹, Bingchen Song², Yuanyuan Zhang³, Haolong Yang¹, Ke Xu¹

1 Beihang University

2 Beijing Institute of Astronautical Systems Engineering

3 North China Institute of Computing Technology

Motivation

Mobile services give rise to numerous spatiotemporal applications.

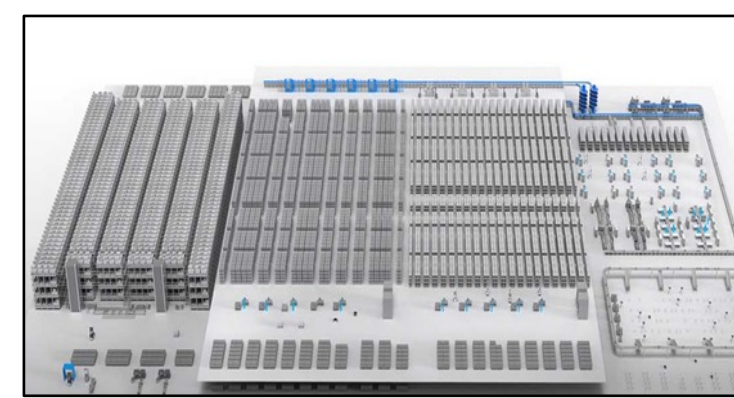
Algorithms determine the operational efficiency in these domains



RoboTaxi



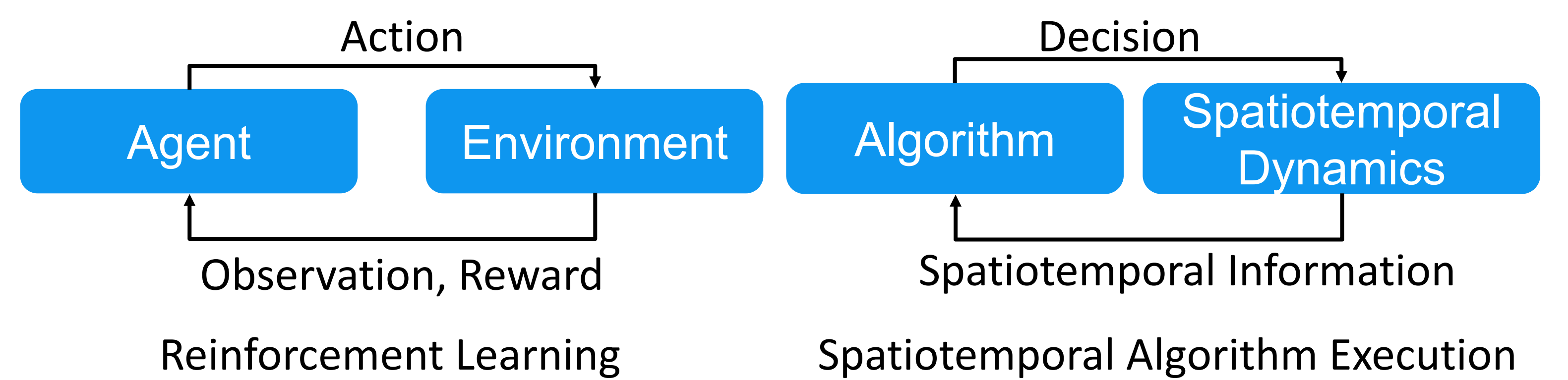
Loading task scheduling



Warehouse robots

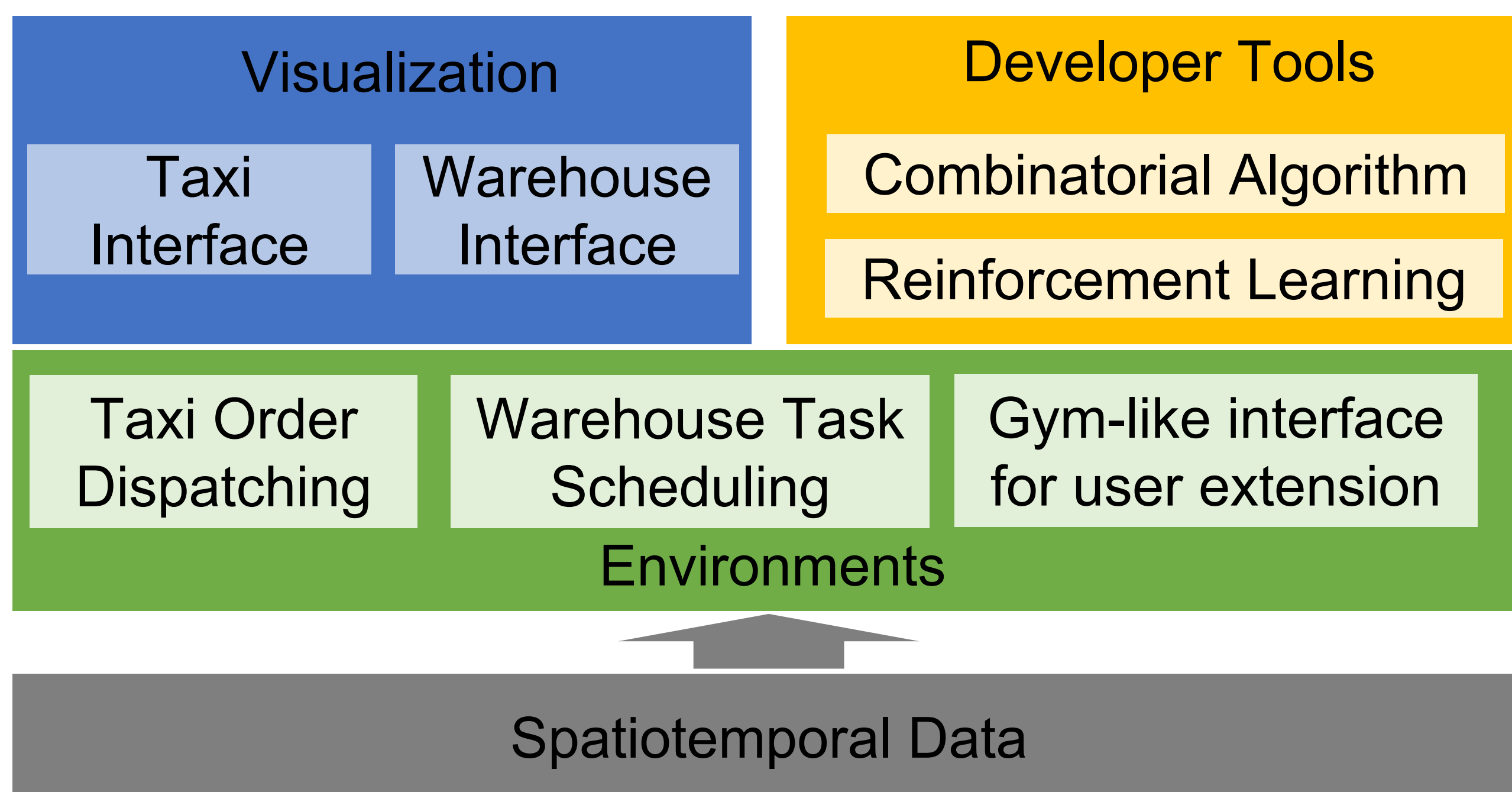
The growing body of spatiotemporal data has spurred a surge in reinforcement learning

No specially designed simulation system for spatiotemporal algorithms



Data-driven Spatiotemporal Simulator (DSS)

Framework



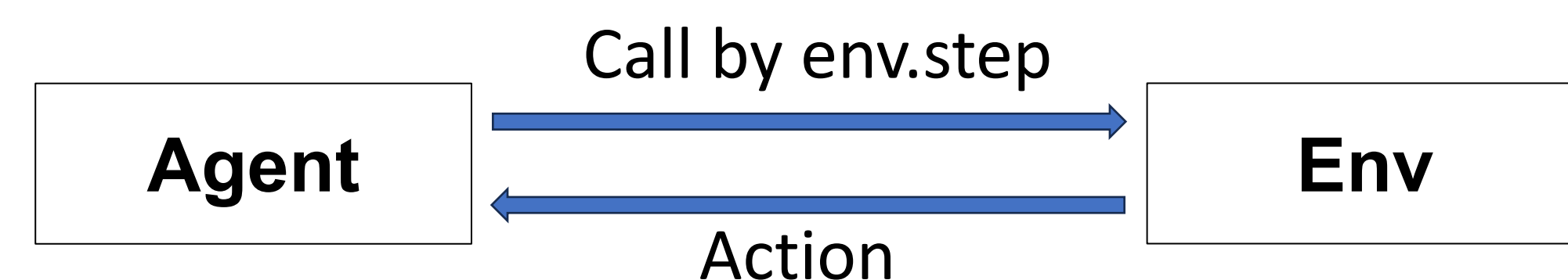
Supported Algorithms

Reinforcement Learning
PPO/REINFORCE/DQN/DDPG/LAF/C51/...
Prioritized Experience Replay

Combinatorial Optimization
Hungarian/Kuhn-Munkres/Dijkstra's/A*/...

Two Modes

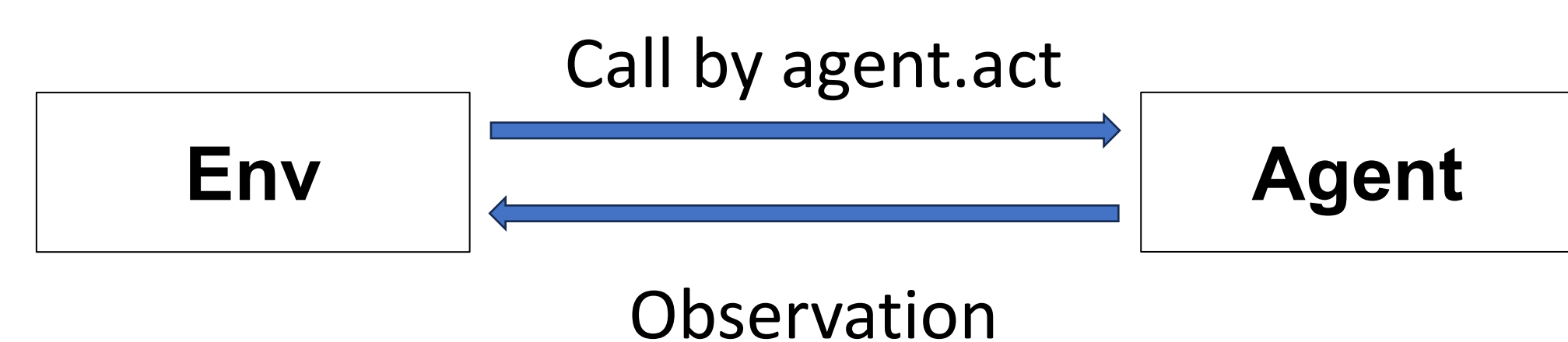
Agent-central mode (Gym like)



```
# initialize an env
env = Maker.make("taxi")
agent = Agent()
matching_action = None
repo_action = None
observation, done = env.reset(date=20161106, start_time=14400, finish_time=20000, False)
while not done:
    dispatch_observ = observation.get('dispatch_observ', None)
    matching_action = agent.action(dispatch_observ) if dispatch_observ else None
    observation, reward, done, info = env.step(dict(matching=matching_action, repo=repo_action))
```

```
(dss) dingyuan@bda-group:~/DSS/dss/envs$ export DSS_DIR=/DSS/ && python -m tests.test_env_taxi
loading from day20161106...
finish loading, begin simulate
04:10: 0/ 920 drivers, 216 orders, r_ana: 0.484, r_comp: 0.459, match time 0.002, repo time 0.000, rewards: 888.040, total rewards: 888.040
04:20: 747/ 920 drivers, 214 orders, r_ana: 0.458, r_comp: 0.441, match time 0.004, repo time 0.000, rewards: 773.020, total rewards: 1661.060
04:30: 669/ 920 drivers, 229 orders, r_ana: 0.447, r_comp: 0.413, match time 0.006, repo time 0.000, rewards: 913.450, total rewards: 2574.510
04:40: 662/ 920 drivers, 164 orders, r_ana: 0.335, r_comp: 0.311, match time 0.007, repo time 0.000, rewards: 639.400, total rewards: 3213.970
04:50: 707/ 920 drivers, 166 orders, r_ana: 0.372, r_comp: 0.347, match time 0.009, repo time 0.000, rewards: 748.700, total rewards: 3962.670
05:00: 713/ 920 drivers, 166 orders, r_ana: 0.342, r_comp: 0.311, match time 0.010, repo time 0.000, rewards: 638.380, total rewards: 4601.050
05:10: 723/ 920 drivers, 155 orders, r_ana: 0.309, r_comp: 0.287, match time 0.011, repo time 0.000, rewards: 568.060, total rewards: 5169.070
05:20: 724/ 920 drivers, 140 orders, r_ana: 0.297, r_comp: 0.252, match time 0.013, repo time 0.000, rewards: 571.480, total rewards: 5740.550
05:30: 735/ 917 drivers, 141 orders, r_ana: 0.313, r_comp: 0.302, match time 0.014, repo time 0.000, rewards: 696.540, total rewards: 6437.090
```

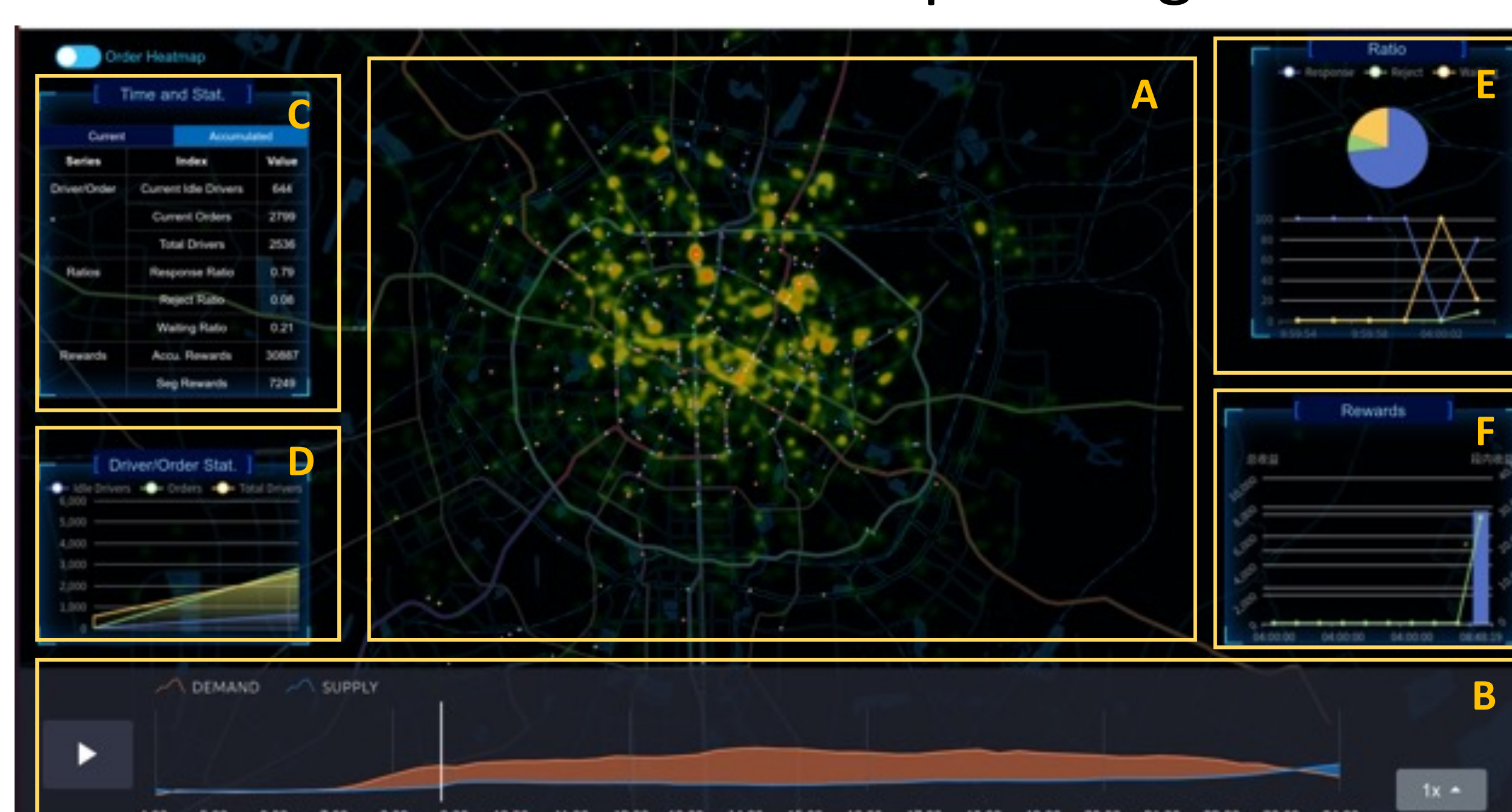
Environment-central mode



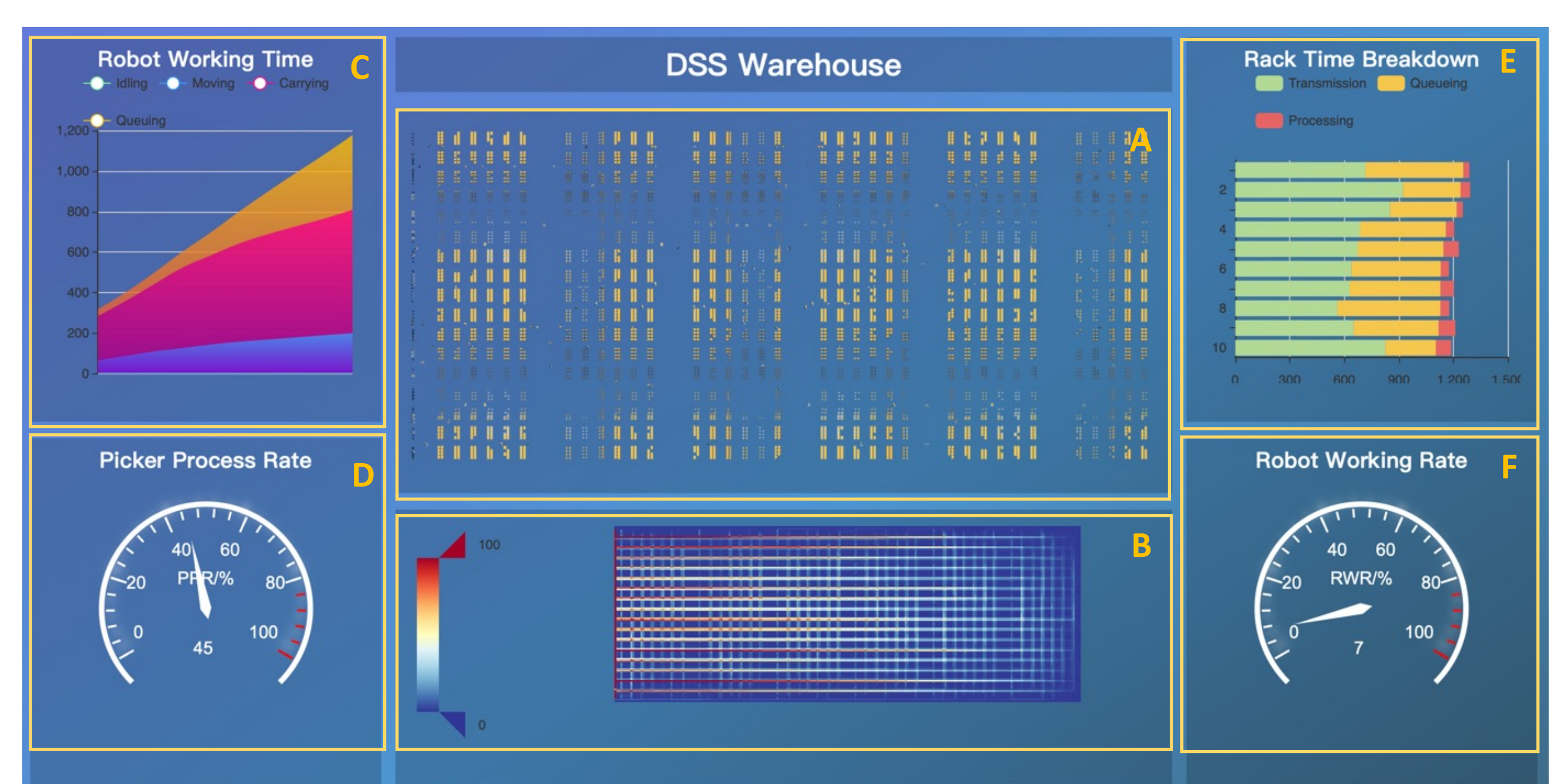
```
(dss) dingyuan@bda-group:~/DSS/dss/envs$ export DSS_DIR=/DSS/ && java -jar engines/warehouse-1.0-SNAPSHOT.jar
23x104 ground, with 4896 racks, 408 robots and 68 pickers
assigned 104 tasks, Makespan: 3, PWR: 0.000, RWR: 0.000
assigned 240 tasks, Makespan: 6, PWR: 0.000, RWR: 0.000
assigned 366 tasks, Makespan: 9, PWR: 0.000, RWR: 0.000
assigned 468 tasks, Makespan: 179, PWR: 0.428, RWR: 0.070
assigned 569 tasks, Makespan: 277, PWR: 0.574, RWR: 0.095
assigned 671 tasks, Makespan: 344, PWR: 0.616, RWR: 0.102
assigned 771 tasks, Makespan: 434, PWR: 0.584, RWR: 0.097
assigned 872 tasks, Makespan: 498, PWR: 0.553, RWR: 0.091
assigned 972 tasks, Makespan: 562, PWR: 0.526, RWR: 0.087
assigned 1072 tasks, Makespan: 628, PWR: 0.508, RWR: 0.084
assigned 1175 tasks, Makespan: 765, PWR: 0.491, RWR: 0.081
assigned 1276 tasks, Makespan: 927, PWR: 0.473, RWR: 0.078
```

System Snapshot

Taxi Order Dispatching



Warehouse Task scheduling



Source code: <https://github.com/dingyuan-shi/spatiotemporal-simulator-rl>