東京大学　関本研究室 / Sekimoto Lab. IIS, the University of Tokyo.

Decision-Making System for Road-Recovery Considering Human Mobility by Applying Deep Q-Learning

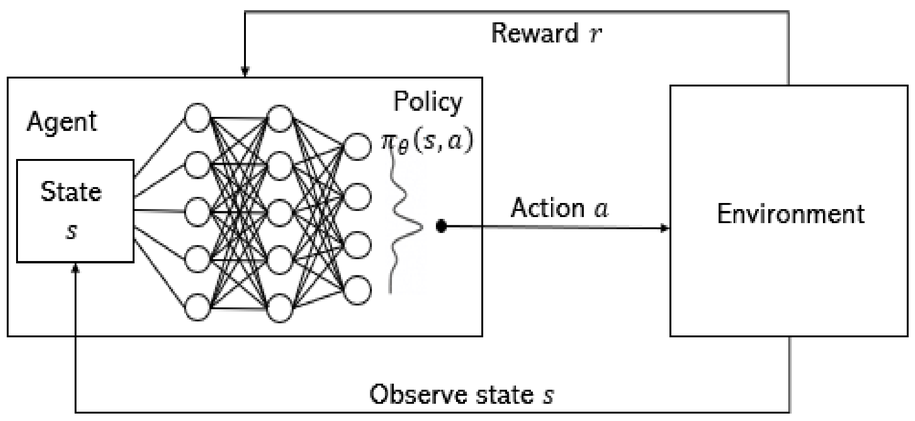
JOO Soohyun, Yoshiki Ogawa, Yoshihide Sekimoto

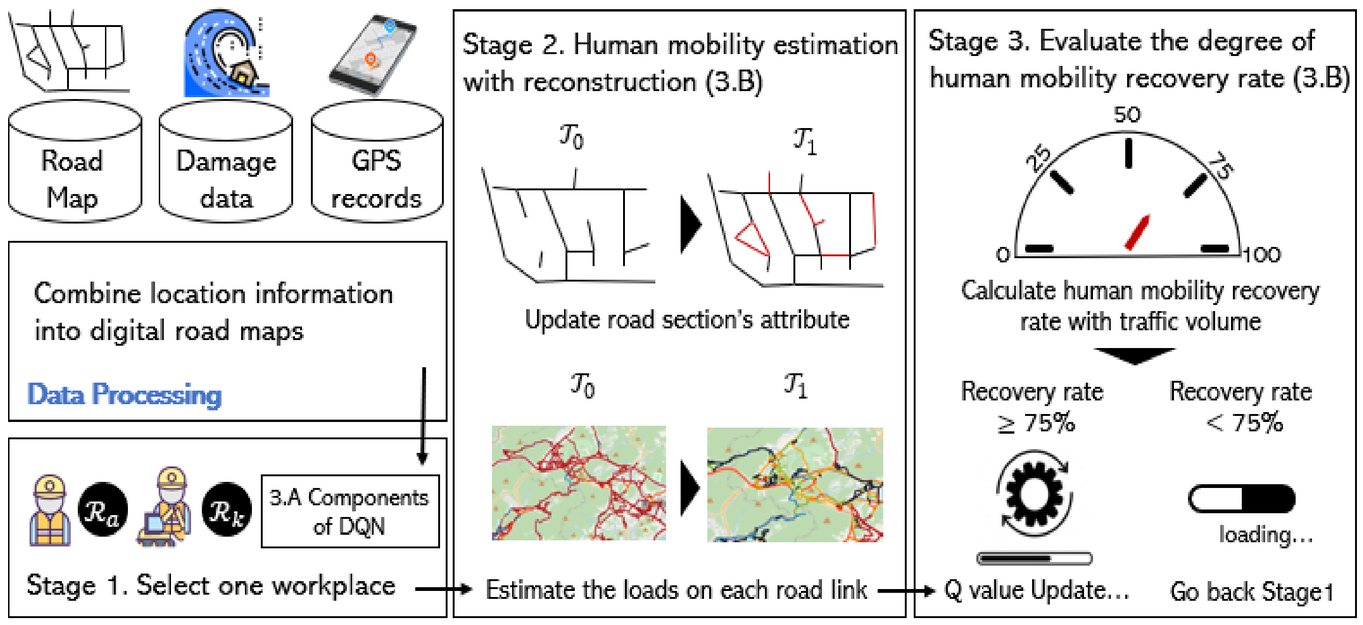
Background

Western Japan experienced heavy, record-breaking rain from June 28 to July 8, 2018, causing approximately 600 road sections to be closed in Hiroshima and Okayama Prefecture. The government wanted the rapid recovery of disrupted human mobility, however, restoring roads frequently used by citizens was delayed for a week after this flooding. There are three assignments for efficient disaster road management plan: 1) the lack of priori knowledge, 2) the absence of evaluation indicators, and 3) the non-consideration of human mobility changes in post-disaster. To solve these limitations, we develop data-driven reinforcement learning modeling to human movement in recovery operation by utilizing mobile phone GPS data.

Proposed System

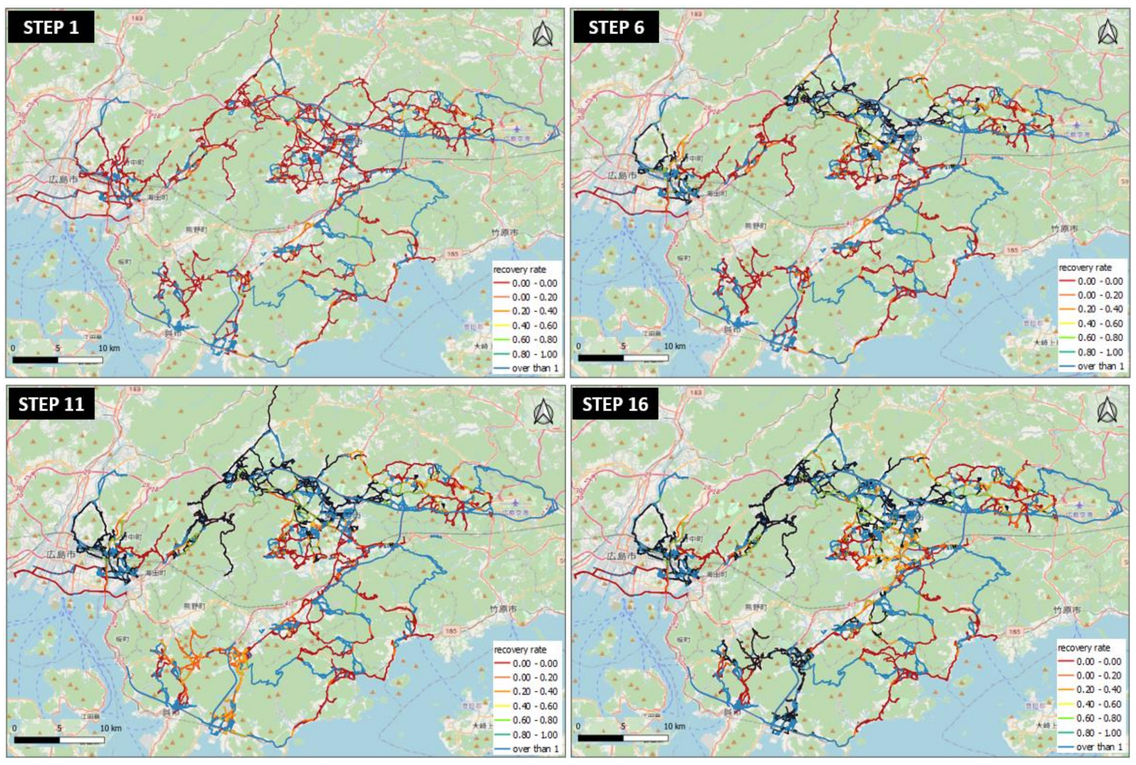
Deep Reinforcement learning generates insights and finds optimal answer via the interaction with the ever-changing environment.

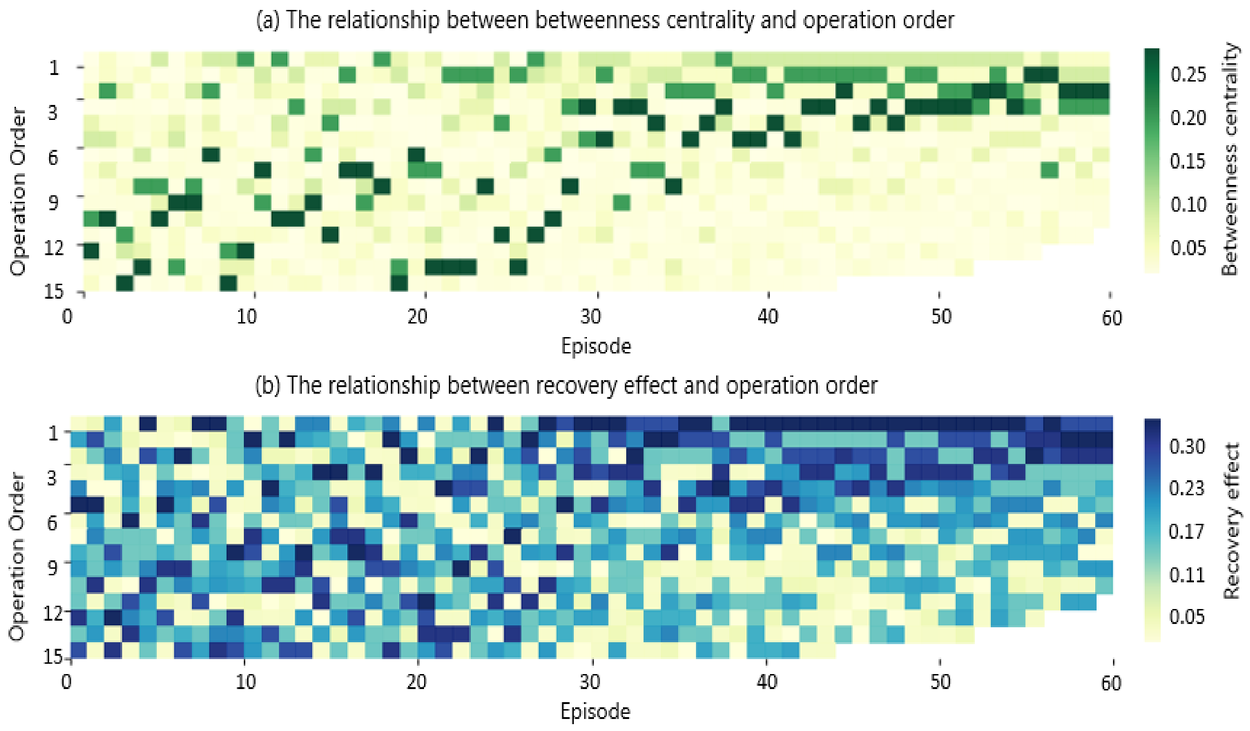




We utilized multi-agent RL to build the mathematical model of multi-locational road recovery plan from collected data of human mobility, recovery situation.

Analysis Result





- As the left figure shown, we could confirm that the damage road situated near the urban district is first to reconstruct.
- Providing dynamic traffic information as input data, the agents tend to select central roads in transport network preferentially.
- The work crews could recognize which damage road has high effect on human mobility recovery.

Summary

We identified the optimal management for damage roads located in three municipal regions (Hiroshima, Higashi-Hiroshima, Soja). The final human mobility recovery rate is approximately 20% better on average than the lowest recovery rate when working randomly. Approximately 1,000 kinds of O-Ds were used to estimate human mobility with reconstruction. The agents in this model would choose preferentially the disrupted road with higher recovery effect and recognize the centrality of each target receiving the information of the degree of recovery and changing road usage situation. The number of cases in which the agent should explore to obtain optimal policy is $2.32 * 10^{52}$. It takes about 3 hours to determine the optimal recovery plan.