Privacy Preserving Planning in Stochastic Environments

Guy Shani, Roni Stern and Tommy Hefner

Extending Collaborative Privacy Preserving Planning by formulating an algorithm to solve domains with stochastic effects

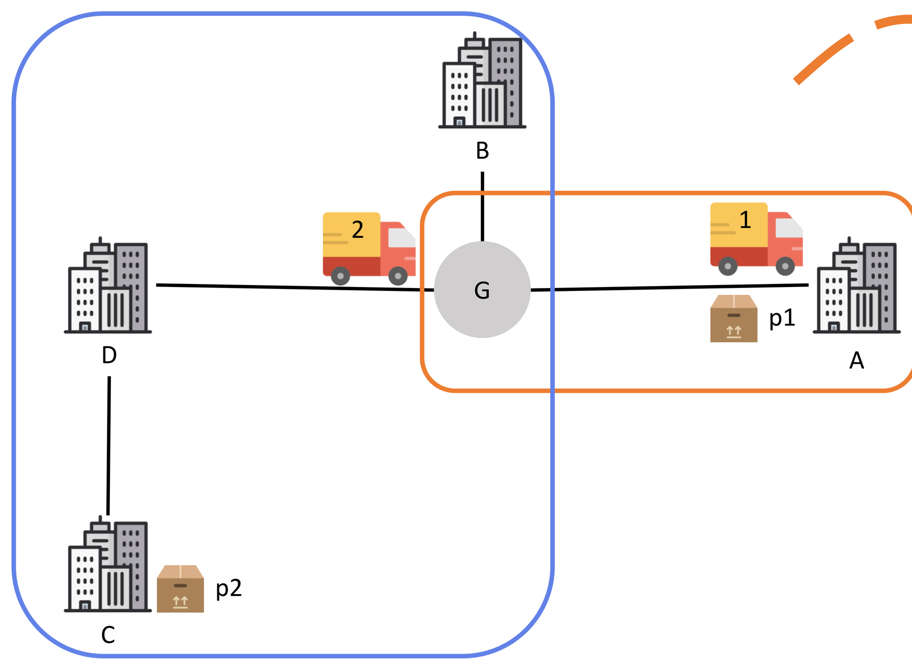
CPPP

Collaborative Privacy Preserving Planning – CPPP

Public State –
- P1 at city A
- P2 at city B

Private State –
- T1 at city A
- T2 at city C

Goal – get package P1 to city B and package P2 to city D without revealing private information.



DRTDP

Distributed Real Time Dynamic Programming – DRTDP

Running simulated trajectories from the initial state to the goal state. Collaborating by passing the trajectory from one agent to another according to the different Q values.

Updating Q values with the Bellman Equation using valued received from other agents using a message passing mechanism.

Message Types:
- Q value request – request Q value of a state for Bellman Update
- Q value answer – answer a Q value request
- Trajectory – pass a trajectory to another agent by sending the state from which it should advance

Every iteration of trajectory advancement requires message passing in order to perform the Bellman Update. The algorithm runs identical trajectories to normal RTDP on the joint problem thus maintaining the same convergence properties.

PS-RTDP

Public Synchronization Real Time Dynamic Programming – PS-RTDP

An approximation of DRTDP, running simulated trajectories from the initial state to the goal state. PS-RTDP differentiates between actions that change the public state and actions that only change the private state of some agent.

When a public change occurs, a DRTDP update is done where messages are sent and other agent's Q values might be used. When only a private change occurs, a local Bellman Update is performed using only the acting agents Q values.

Passing messages on public changes only, lead to significant reduction in the amount of messages and effects the algorithm run time as well. There is an issue caused by this as well.

Private Cycles – since a trajectory passes only on public changes, what happens if an agent got the trajectory but has only private actions to perform? A private loop!

A cycle detection mechanism is required in order to avoid getting stuck in private loops. If a cycle is detected, PS-RTDP restarts the trajectory from the initial state.

If the cycle detection mechanism is too sensitive, too many restarts will lead to very slow run times.

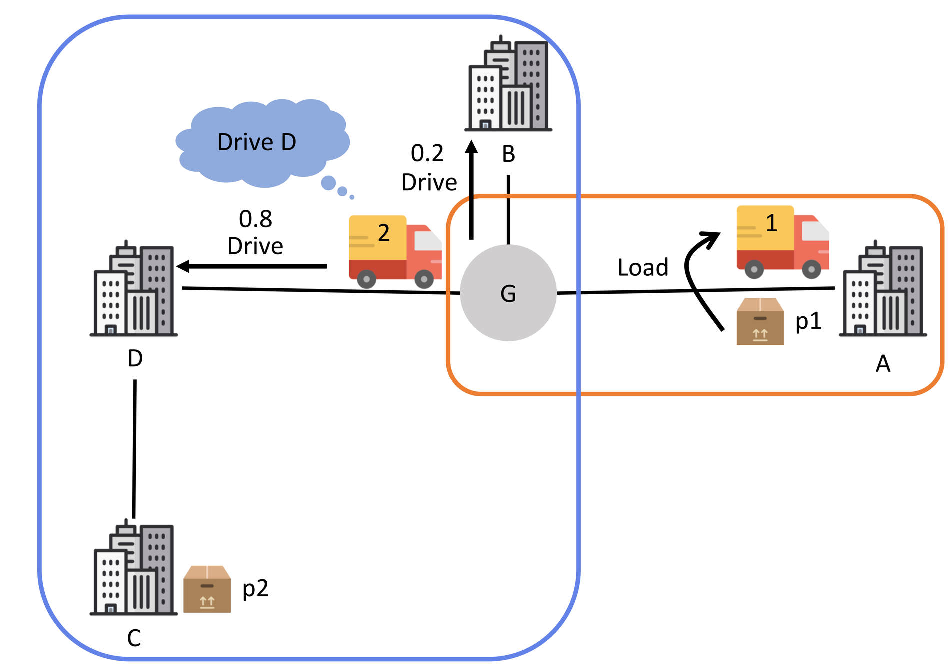
If the cycle detection mechanism is not sensitive enough, precious time will be wasted until a private loop is detected.

SCPPP

Stochastic Collaborative Privacy Preserving Planning – SCPPP

An SCPPP problem is a tuple $< P, \{A_i\}_{i=1}^{k}, tr, I, G, C >$

- P – set of propositions
- S – set of CPPP states made of all possible assignments to P
- A_i – set of actions for agent i, k = number of agents
- tr – transition function $tr\colon SxAxS \rightarrow \mathbb{R}$
- I – initial state
- G – goal state
- C – cost function C: $SxA \rightarrow \mathbb{R}$



State

Public:
- P1 at A

Private:
- T1 at A
- Index of (T2 at G)

Actions $\{A_i\}_{i=1}^{k}$

T1:
- Load P1
- Drive G

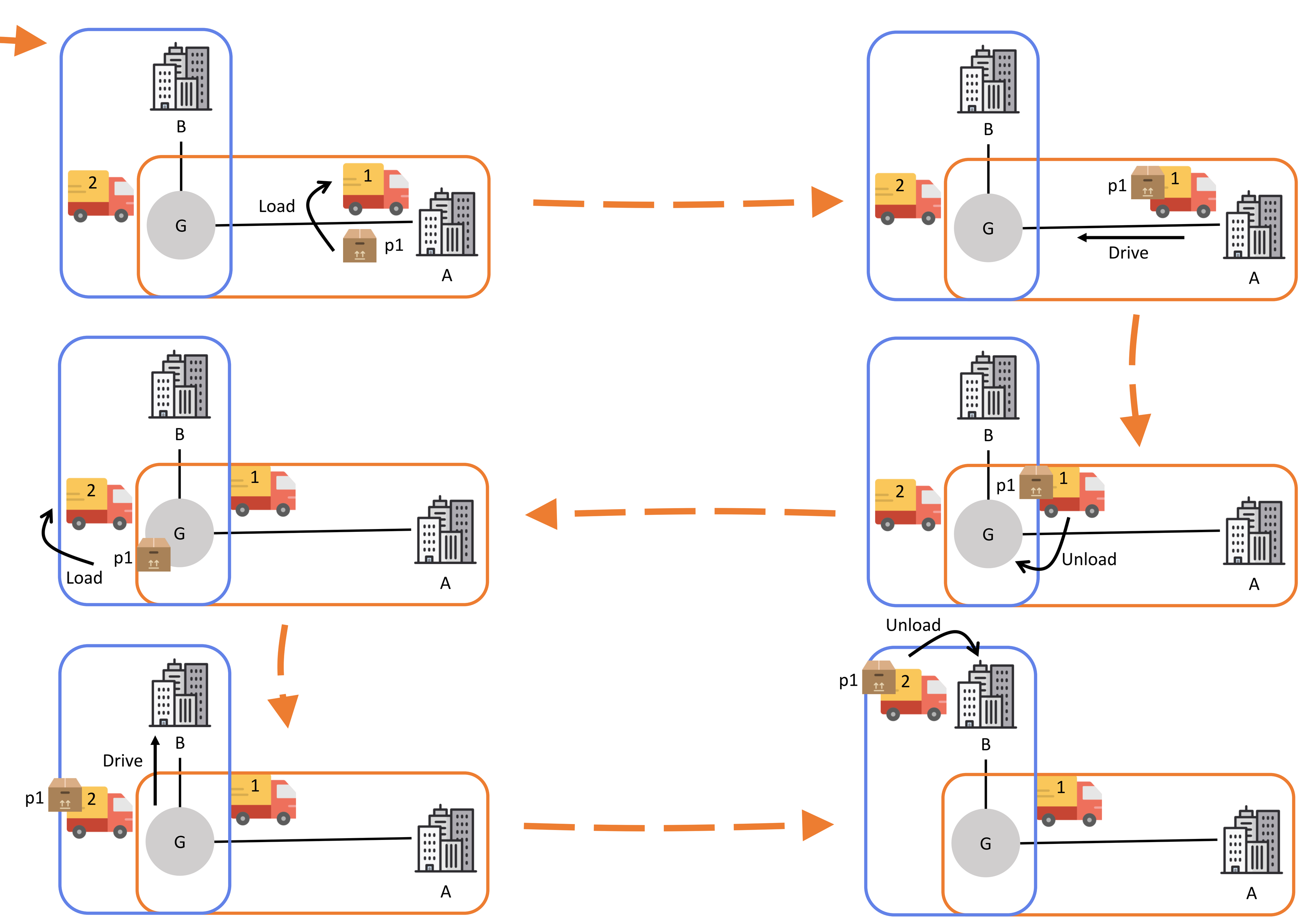
T2:
- Drive B
- Drive D

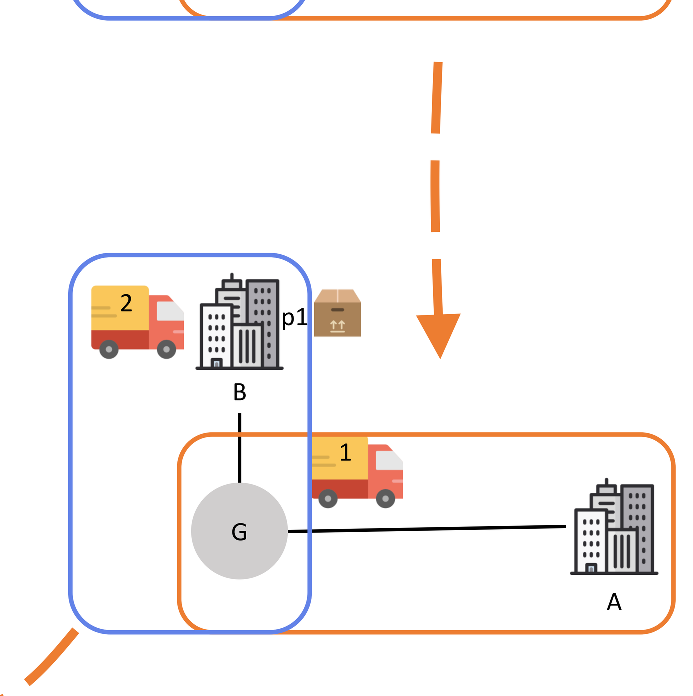
Effects

T1 Load P1:
- P1 loaded 0.9
- P1 still in A 0.1

T2 Drive B:
- T2 in B 0.2
- T2 in D 0.8

Trajectory Example



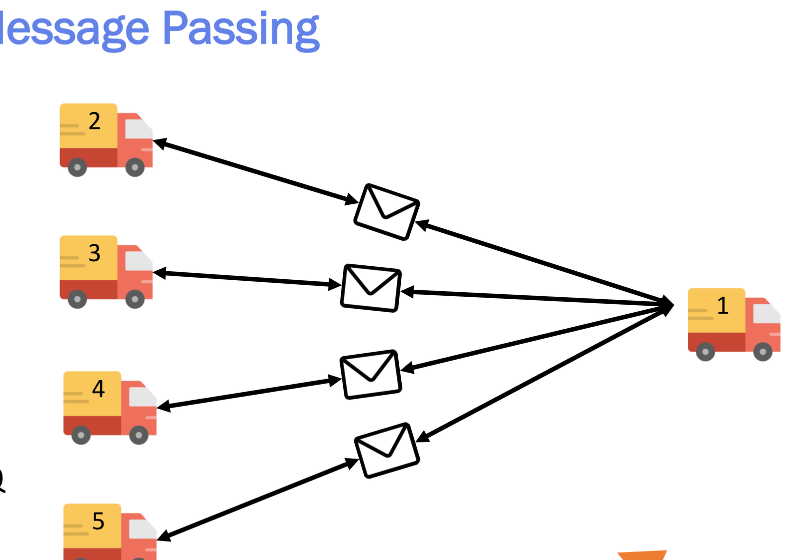


Bellman Update Message Passing

T1 broadcasts a Q value request message for a Bellman update.

The rest of the agents answer with their Q values.

Then T1 can update its Q value using the Bellman Equation.



$$Q_i(s,a) = C(s,a) + \sum_{s'} tr(s,a,s')V_i(s') \qquad V_i(s) \leftarrow min_{a'} Q_i(s,a')$$

Experiments

Three domains adapted from CPPP.

Experiments ran on a Google Cloud VM with 2GHz CPU with 2 cores.

The algorithms were implemented in Python.

Blocksworld – robots stacking blocks on a table. Different robots can interact with different blocks. Collaboration requirement depends on problem setup.

Little private information, deep goals.

Depot – drivers moving crates between depots and warehouses, lifts loading, unloading and ordering crates in the depots / warehouses.

Collaboration always required, medium amount of private information.

Logistics – trucks and airplanes moving packages between cities.

Collaboration depends on problem setup. Most information is private. Deep goals.

Domain	# Actions	Best Cost	Expansions ×10^4	Messages ×10^4	# trajectories + restarts	Total Time (sec)
blocks-5-2	422	(**11.98** / 12.36)	(**5.247** / 5.412)	(30.434 / **9.366**)	(150 / 380 + 691)	(35.3 / **29.8**)
blocks-6-2	746	(**14.8** / 16.5)	(**68.322** / 68.977)	(399.718 / **117.096**)	(900 / 2140 + 7057)	(497.8 / **427.7**)
depot-2-5	245	(11.4 / **11.14**)	(**4.342** / 5.050)	(89.343 / **23.417**)	(350 / 170 + 1389)	(55.7 / **23.9**)
depot-3-5	407	(16.78 / 16.78)	(**55.045** / 59.236)	(1136.297 / **289.422**)	(2330 / 740 + 10437)	(715.4 / **344.8**)
logistics-2-5	203	(27.0 / **21.58**)	(122.934 / **40.523**)	(2705.937 / **103.567**)	(2134 / 220 + 14162)	(848.5 / **81.7**)
logistics-3-4	104	(**26.8** / 27.22)	(152.786 / **109.649**)	(2570.205 / **250.775**)	(8010 / 770 + 27091)	(720.6 / **186.0**)