QNI Summer Brief

Jason Jabbour
Systems Engineer
2022

JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

Roadmap

- Problem Description
- Methodology Exploration
- Reinforcement Learning (RL) Problem Formulation
- Setting up the RL Toolchain
- Policy Demonstration
- Future Work

Roadmap

- **Problem Description**
- Methodology Exploration
- Reinforcement Learning (RL) Problem Formulation
- Setting up the RL Toolchain
- Policy Demonstration
- Future Work

Problem Description

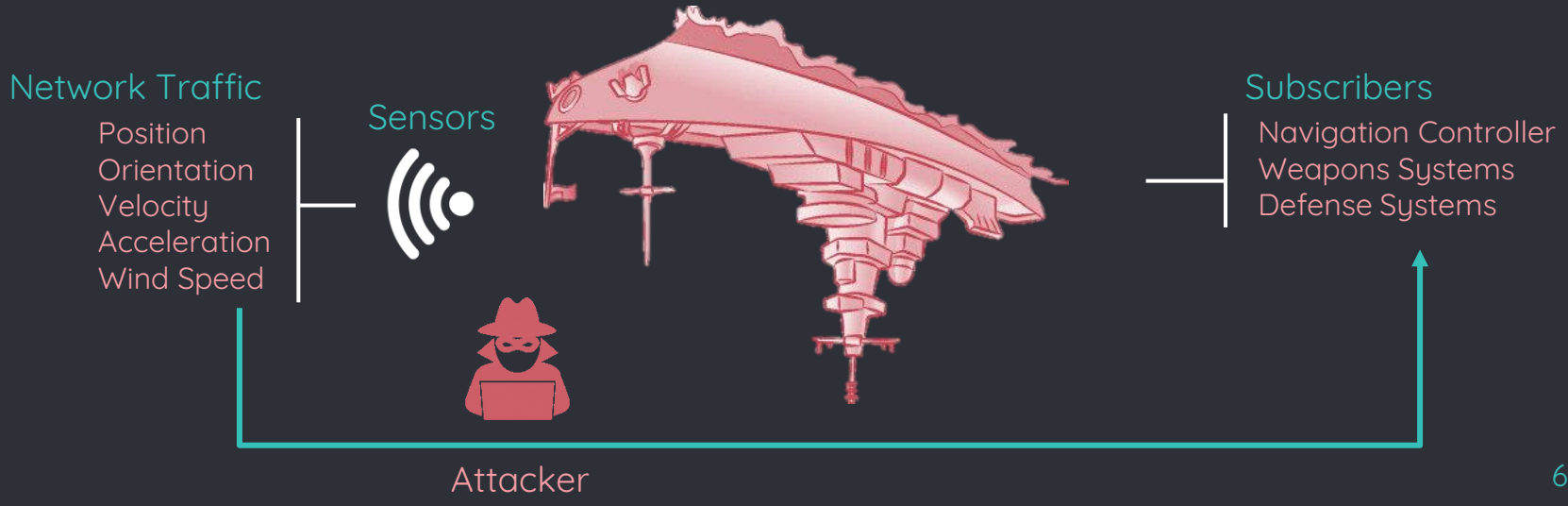
- **Problem**: Network traffic on board US Navy Ships are vulnerable to malicious attacks
- Why is network traffic important?

Problem Description

- **Problem**: Network traffic on board US Navy Ships are vulnerable to malicious attacks
- Why is network traffic important?

Problem Description

- **Problem**: Network traffic on board US Navy Ships are vulnerable to malicious attacks
- Why is network traffic important?



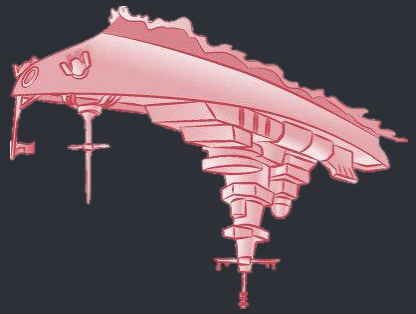
Problem Description

- **Problem**: Network traffic on board US Navy Ships are vulnerable to malicious attacks
- Why is network traffic important?
- **Objective**: Detect malicious attacks to the network traffic data and trigger subsequent alerts.

Expected Ship State

≠

Data should not trusted.
Send Alert!



Network Traffic Data

Roadmap

- Problem Description
- **Methodology Exploration**
- Reinforcement Learning (RL) Problem Formulation
- Setting up the RL Toolchain
- Policy Demonstration
- Future Work

Methodology Exploration

- **Objective**: Detect malicious attacks to the network traffic data and trigger subsequent alerts.
- Utilize machine learning to learn the behavior of a ship

Methodology Exploration

- **Objective**: Detect malicious attacks to the network traffic data and trigger subsequent alerts.
- Utilize machine learning to learn the behavior of a ship

Methods

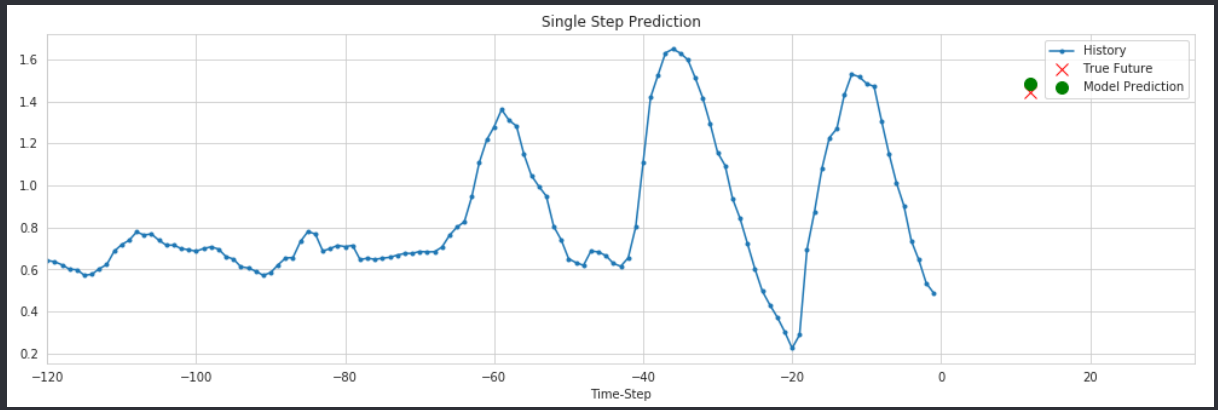
1. **Long Short Term Memory (LSTM) Neural Networks**

2. **Reinforcement Learning**

Methodology Exploration

Methods

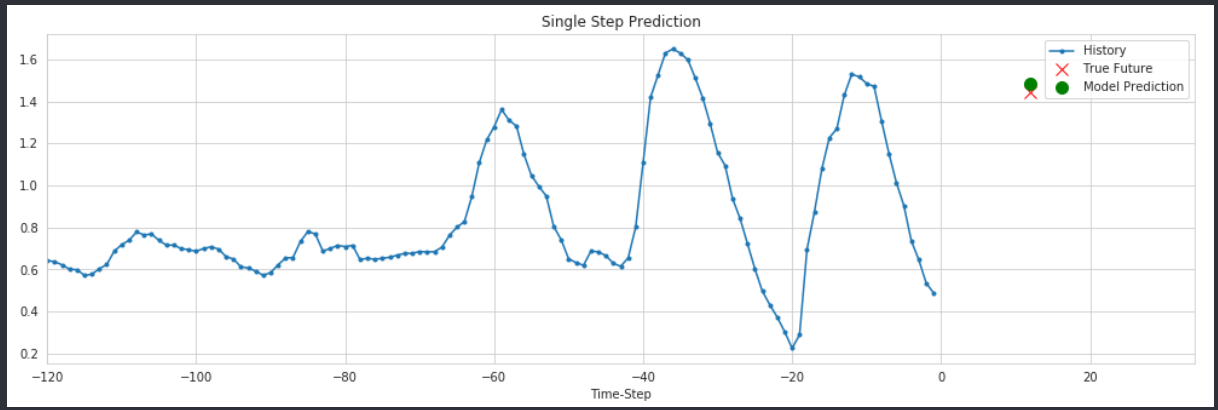
1. **Long Short Term Memory (LSTM) Neural Networks**



Methodology Exploration

Methods

1. **Long Short Term Memory (LSTM) Neural Networks**

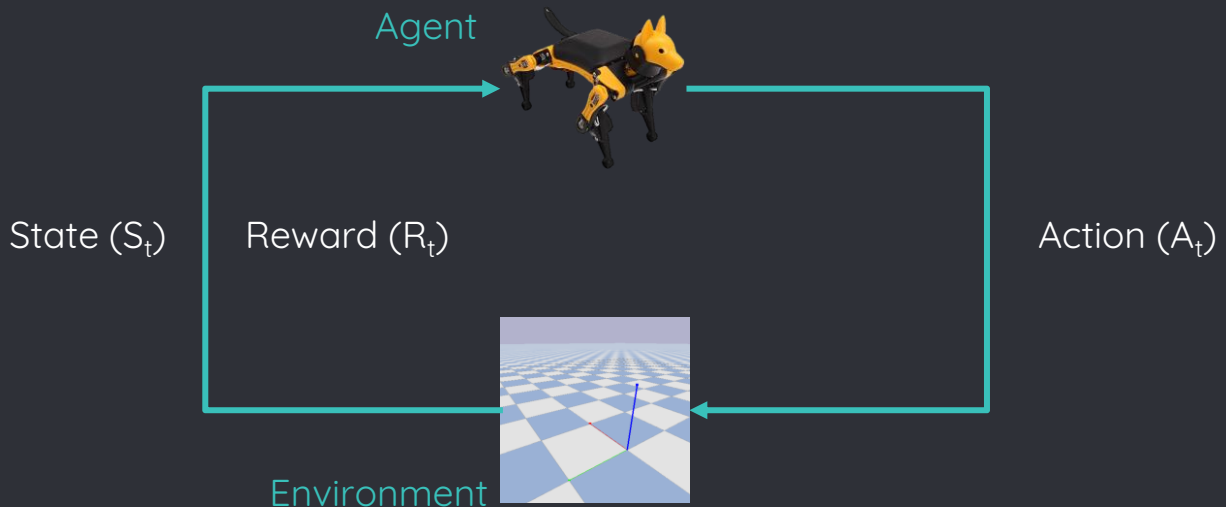


Disadvantage: Does not detect attacks more sophisticated than altering data magnitude

Methodology Exploration

Methods

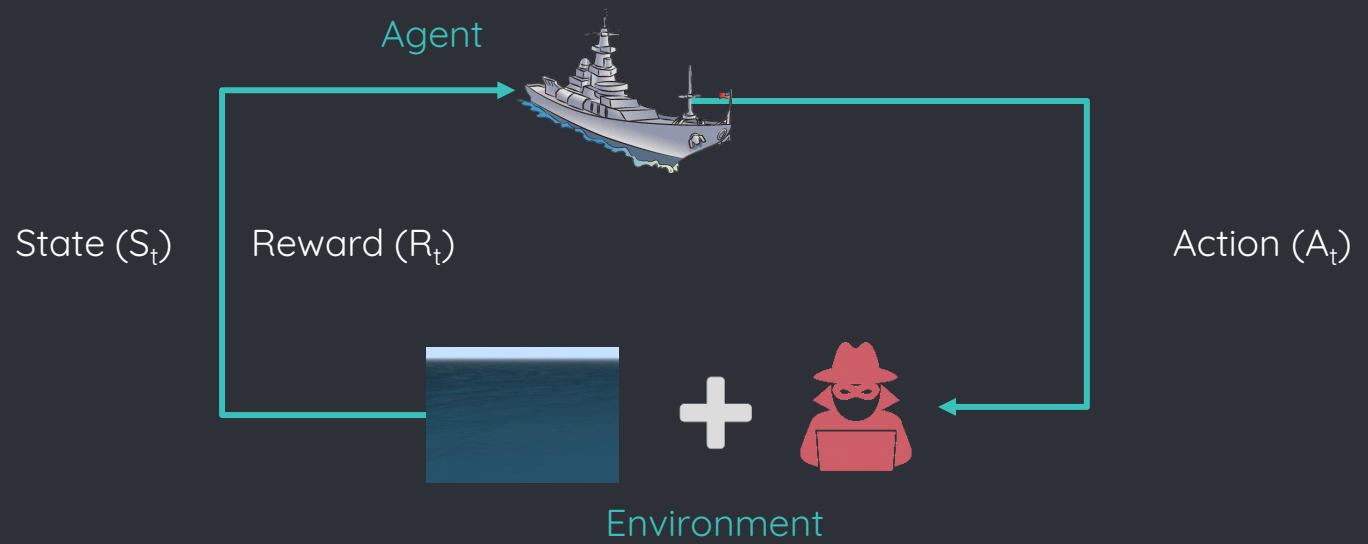
1. Long Short Term Memory (LSTM) Neural Networks
2. **Reinforcement Learning**



Methodology Exploration

Methods

1. Long Short Term Memory (LSTM) Neural Networks
2. **Reinforcement Learning**



Roadmap

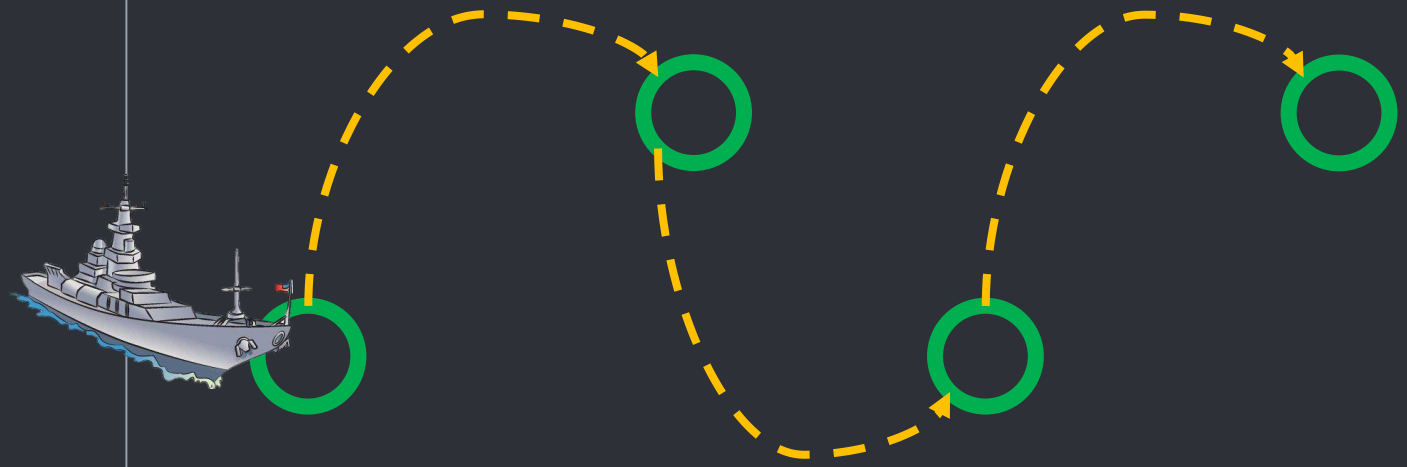
- Problem Description
- Methodology Exploration
- **Reinforcement Learning (RL) Problem Formulation**
- Setting up the RL Toolchain
- Policy Demonstration
- Future Work

RL Problem Formulation

- **Per Episode:**
  - Randomly generate waypoints
  - Use a PID controller to navigate to waypoints
  - Randomly set environment factors

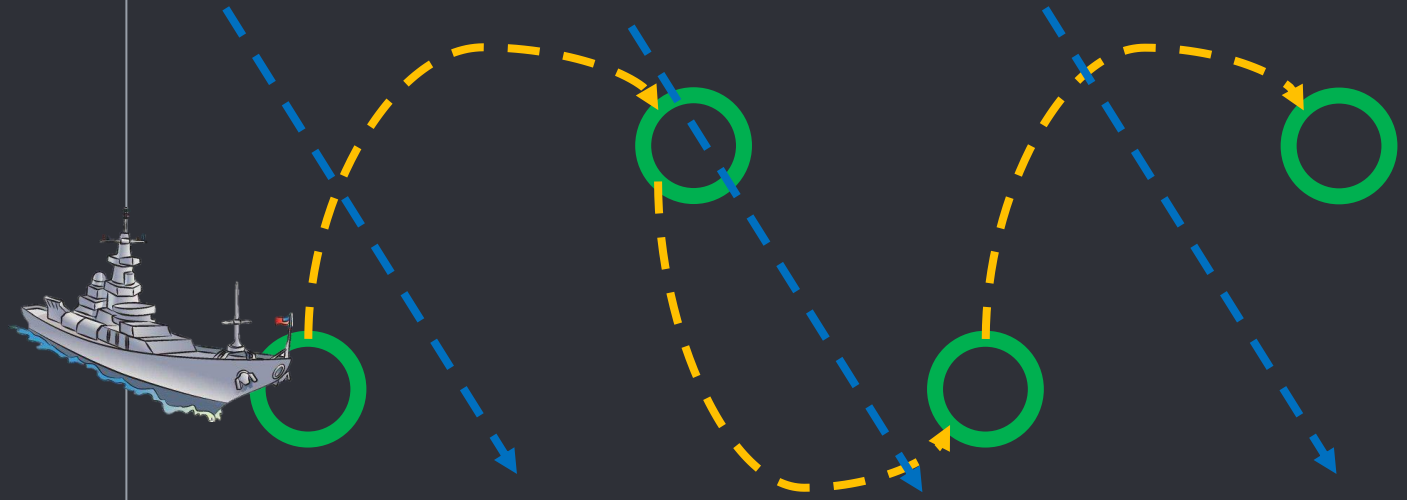
RL Problem Formulation

- **Per Episode:**
  - Randomly generate waypoints
  - Use a PID controller to navigate to waypoints
  - Randomly set environment factors



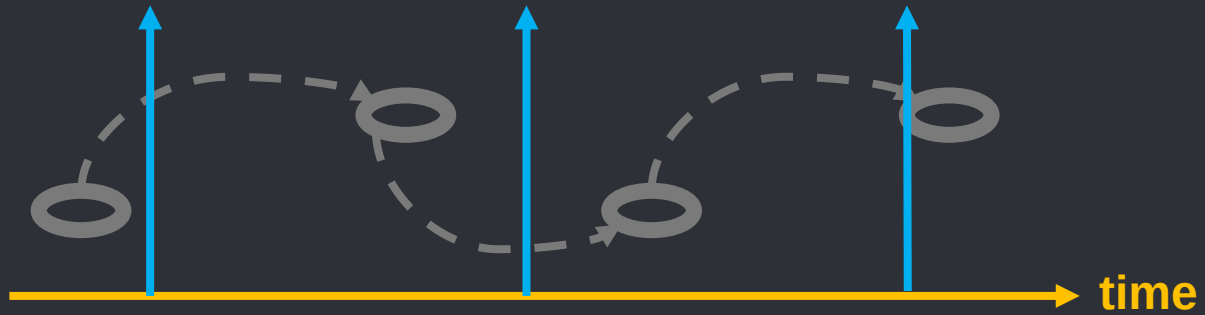
RL Problem Formulation

- **Per Episode:**
  - Randomly generate waypoints
  - Use a PID controller to navigate to waypoints
  - Randomly set environment factors



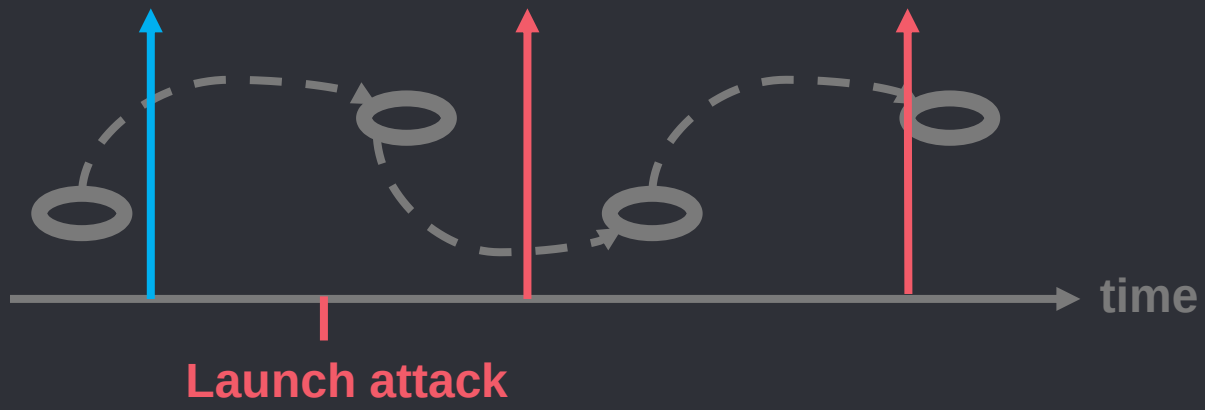
RL Problem Formulation

- **Per Step:**
  - **Collect an observation**
  - **Wrap observation with an attack module**



RL Problem Formulation

- Naïve Attack Module
  - Randomly selects timestep to launch attack
  - Randomly selects observations to attack
  - Randomly selects perturbation amount
  - Continues to attack observation until agent detects attack



RL Problem Formulation

- **Observational Space (O)**

$$
\begin{aligned}
&O = [\\
&[position + n\ history],\\
&[orientation + n\ history],\\
&[position\ rate + n\ history],\\
&[orientation\ rates + n\ history],\\
&[environment\ info + n\ history],\\
&[engine\ info + n\ history],\\
&[next\ waypoint]\\
&]
\end{aligned}
$$

$$
\begin{aligned}
&where:\\
&n=5
\end{aligned}
$$

RL Problem Formulation

- **Action Space (A)**

$A = [A_1, A_2]$

$A_1 = \forall [(o \in \boldsymbol{O}) \wedge \sim (o \in \{sep\})]: 0 \leq a \leq 1$
$A_2 = 0 \leq a \leq 1$

$\boldsymbol{where}$:

$$sep \coloneqq simulation\ exclusive\ features$$

RL Problem Formulation

- **Reward Function (R)**

$$\boldsymbol{R} = \sum_{\boldsymbol{a} \in \boldsymbol{A_1}} \boldsymbol{r}\,(\boldsymbol{a}) + \boldsymbol{b}$$

$\boldsymbol{where}$:

$$r(a) = \begin{cases} [r = a \mid o = true] \\ [r = -a \mid o = false] \end{cases}$$

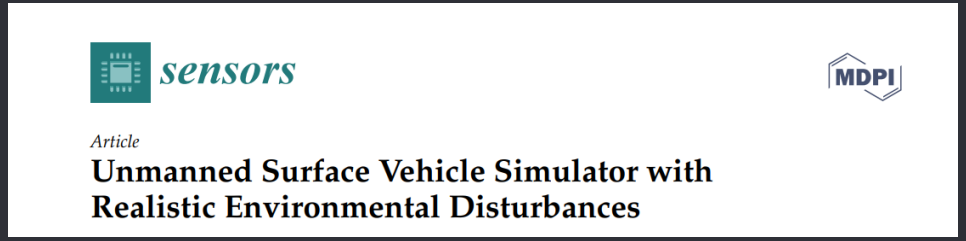
$$b = \begin{cases} [10 \mid (Underattack \wedge (A_2 \geq .8))] \\ [10 \mid (\sim Underattack \wedge (A_2 \leq .2))] \\ [-10 \mid (Underattack \wedge (A_2 \leq .5))] \\ [-10 \mid (\sim Underattack \wedge (A_2 \geq .5))] \end{cases}$$

Roadmap

- Problem Description
- Methodology Exploration
- Reinforcement Learning (RL) Problem Formulation
- **Setting up the RL Toolchain**
- Policy Demonstration
- Future Work

Setting up the RL Toolchain

- **Finding a Surface Vehicle Simulator**

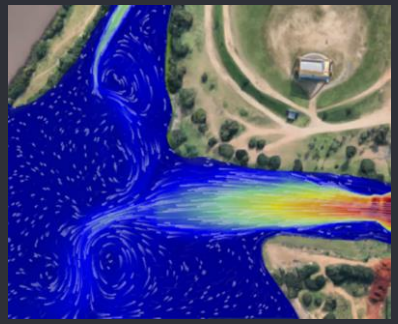


- Benchmarked existing simulators and found a lack of modeling environmental disturbances

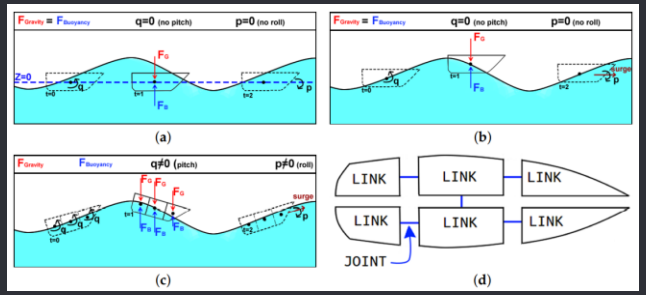
Simulator	Waves	Buoyancy	Water Currents	Wind Currents	Thruster Underwater	Thruster above Water	Foil
UWSim	√	√	×	×	√	×	×
Gazebo	×	×	×	×	√√	√√	×
Freefloating Gazebo	√	√	√	×	√√	√√	×
VREP	√	√	×	×	√	√√	×
RobotX Simulator	√	√√	×	√	√√	√√	×
USVSim	√√	√√	√√	√√	√√	√√	√√

Setting up the RL Toolchain

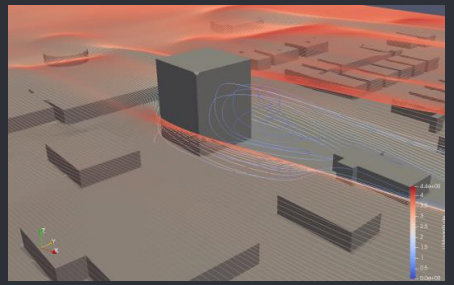
- **ROS based Surface Vehicle Simulator**



Water Currents



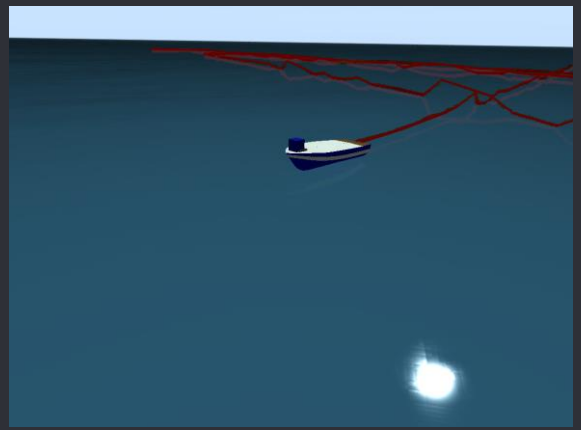
Waves interacting with multiple ship links



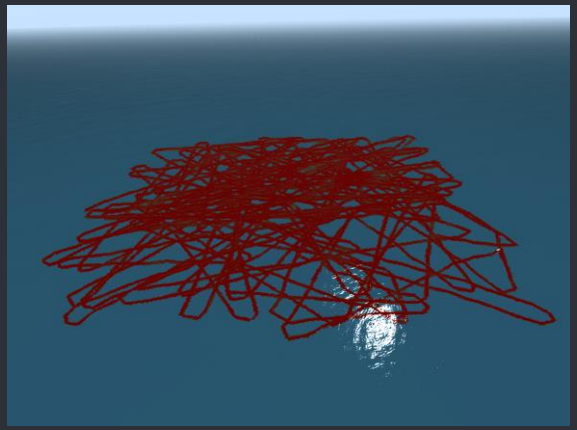
Wind Currents

Setting up the RL Toolchain

- **Extended the ROS based Simulator by developing a node to:**
  - **Publish randomized navigation goals**
  - Collect data by subscribing to relevant topics
  - Set environmental parameters



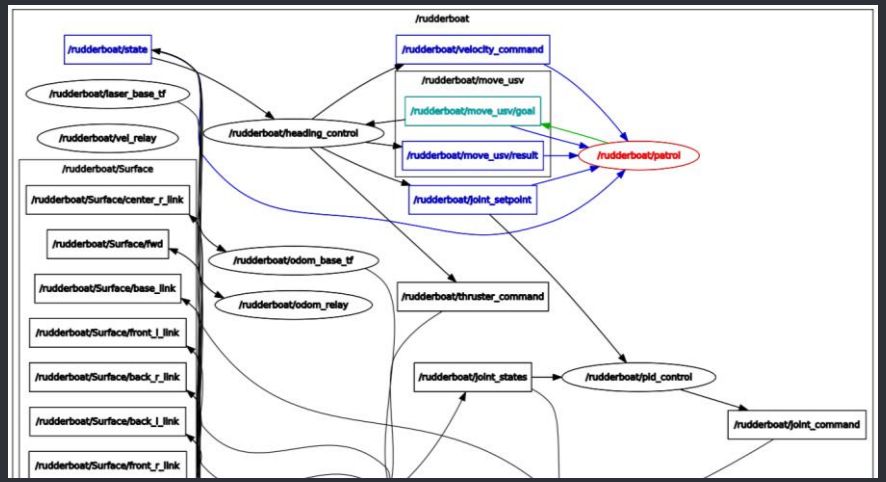
Ship Navigating to Waypoint



Ship Navigating to 300 Waypoints

Setting up the RL Toolchain

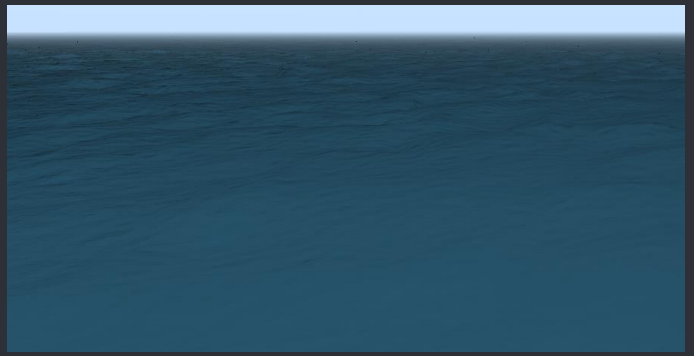
- **Extended the ROS based Simulator by developing a node to:**
  - Publish randomized navigation goals
  - Collect data by subscribing to relevant topics
  - Set environmental parameters



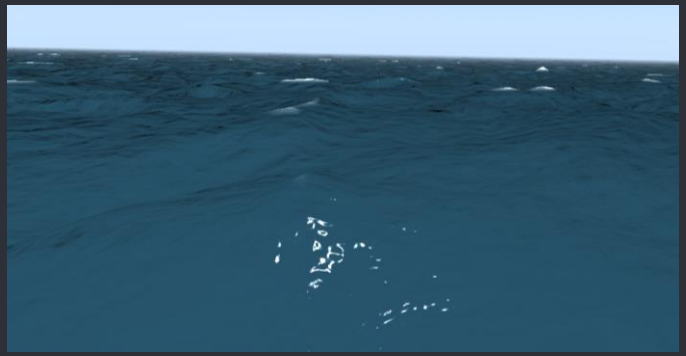
ROS Node Receiving Data and Publishing Waypoints

Setting up the RL Toolchain

- **Extended the ROS based Simulator by developing a node to:**
  - Publish randomized navigation goals
  - Collect data by subscribing to relevant topics
  - **Set environmental parameters**



Calm Water at Low Wind Speed



Rough Waves at High Wind Speed

Setting up the RL Toolchain

- **Toolchain Summary**

Stable-Baselines3

ROS GAZEBO

ROS & GAZEBO

OpenAI Gym

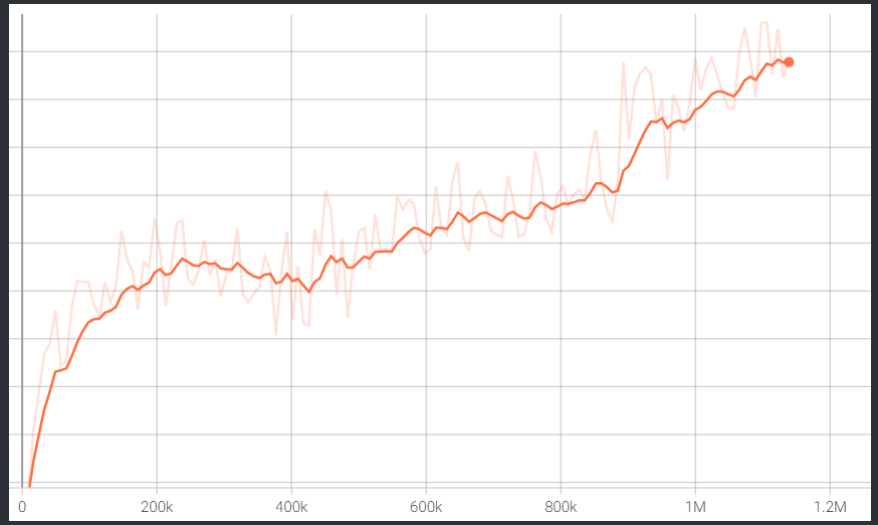
Open AI Gym

Roadmap

- Problem Description
- Methodology Exploration
- Reinforcement Learning (RL) Problem Formulation
- Setting up the RL Toolchain
- **Policy Demonstration**
- Future Work

Policy Demonstration

- **Training a Policy**
  - PPO Algorithm
  - 1 Million Timesteps



Learning Curve

Policy Demonstration

- **Sanity Check**

```
Observation:
[0]     position_x: 263.15
[1]     position_y: 36.92
[2]     position_z: 0.94
[3]     orientation_x: 0.06
[4]     orientation_y: 0.01
[5]     orientation_z: 0.96
[6]     position_rate_x: 1.12
[7]     position_rate_y: -0.01
[8]     position_rate_z: -0.91
[9]     orientation_rate_x: 0.11
[10]    orientation_rate_y: 0.02
[11]    orientation_rate_z: -0.01
[12]    wind_speed: 3.0
[13]    engine_velocity_command: 200.0
[14]    rudder_angle: 0.04
Index to modify: no
Action:
[0]     position_x: 0.8
[1]     position_y: 0.82
[2]     position_z: 0.29
[3]     orientation_x: 0.03
[4]     orientation_y: 0.49
[5]     orientation_z: 0.33
[6]     position_rate_x: 0.97
[7]     position_rate_y: 0.0
[8]     position_rate_z: 0.26
[9]     orientation_rate_x: 0.39
[10]    orientation_rate_y: 0.23
[11]    orientation_rate_z: 0.31
[12]    wind_speed: 0.48
[13]    engine_velocity_command: 0.49
[14]    rudder_angle: 0.87
[15]    Under Attack: 0.0
Next? [Y/n]
```

Observations

Confidence Levels

Benign Traffic

Policy Demonstration

- **Sanity Check**

```
Observation:
[0]     position_x: 263.15
[1]     position_y: 36.92
[2]     position_z: 0.94
[3]     orientation_x: 0.06
[4]     orientation_y: 0.01
[5]     orientation_z: 0.96
[6]     position_rate_x: 1.12
[7]     position_rate_y: -0.01
[8]     position_rate_z: -0.91
[9]     orientation_rate_x: 0.11
[10]    orientation_rate_y: 0.02
[11]    orientation_rate_z: -0.01
[12]    wind_speed: 3.0
[13]    engine_velocity_command: 200.0
[14]    rudder_angle: 0.04
Index to modify: no
Action:
[0]     position_x: 0.8
[1]     position_y: 0.82
[2]     position_z: 0.29
[3]     orientation_x: 0.03
[4]     orientation_y: 0.49
[5]     orientation_z: 0.33
[6]     position_rate_x: 0.97
[7]     position_rate_y: 0.0
[8]     position_rate_z: 0.26
[9]     orientation_rate_x: 0.39
[10]    orientation_rate_y: 0.23
[11]    orientation_rate_z: 0.31
[12]    wind_speed: 0.48
[13]    engine_velocity_command: 0.49
[14]    rudder_angle: 0.87
[15]    Under Attack: 0.0
Next? [Y/n]
```

Benign Traffic

High Confidence Levels

Overall Threat Level

```
Under Attack: 0.0
```

Policy Demonstration

- **Sanity Check**

```
Observation:
[0]     position_x: 263.24
[1]     position_y: 36.87
[2]     position_z: 1.0
[3]     orientation_x: 0.06
[4]     orientation_y: 0.0
[5]     orientation_z: 0.96
[6]     position_rate_x: 1.2
[7]     position_rate_y: 0.02
[8]     position_rate_z: -0.58
[9]     orientation_rate_x: 0.12
[10]    orientation_rate_y: -0.0
[11]    orientation_rate_z: -0.0
[12]    wind_speed: 3.0
[13]    engine_velocity_command: 200.0
[14]    rudder_angle: 0.05
Index to modify: 1
Perturbation: 234102981309812
Action:
[0]     position_x: 0.0
[1]     position_y: 0.0
[2]     position_z: 0.02
[3]     orientation_x: 0.0
[4]     orientation_y: 0.06
[5]     orientation_z: 0.0
[6]     position_rate_x: 0.0
[7]     position_rate_y: 0.0
[8]     position_rate_z: 0.0
[9]     orientation_rate_x: 0.02
[10]    orientation_rate_y: 0.0
[11]    orientation_rate_z: 0.16
[12]    wind_speed: 0.0
[13]    engine_velocity_command: 0.0
[14]    rudder_angle: 0.0
[15]    Under Attack: 0.08
Next? [Y/n]
```

Attacking Y Position

Low Confidence Levels

Overall Threat Level

Under Attack: 0.08

Attack Y Position

Policy Demonstration

- **Sanity Check**

```
Observation:
[0]     position_x: 263.59
[1]     position_y: 36.67
[2]     position_z: 0.9
[3]     orientation_x: 0.05
[4]     orientation_y: -0.01
[5]     orientation_z: 0.96
[6]     position_rate_x: 1.43
[7]     position_rate_y: 0.04
[8]     position_rate_z: -0.89
[9]     orientation_rate_x: 0.05
[10]    orientation_rate_y: -0.07
[11]    orientation_rate_z: 0.0
[12]    wind_speed: 3.0
[13]    engine_velocity_command: 200.0
[14]    rudder_angle: 0.04
Index to modify: 0
Perturbation: 1234021384
Index to modify: 1
Perturbation: 1203981
Index to modify: no
Action:
[0]     position_x: 0.03
[1]     position_y: 0.07
[2]     position_z: 0.02
[3]     orientation_x: 0.04
[4]     orientation_y: 0.07
[5]     orientation_z: 0.05
[6]     position_rate_x: 0.04
[7]     position_rate_y: 0.0
[8]     position_rate_z: 0.0
[9]     orientation_rate_x: 0.16
[10]    orientation_rate_y: 0.12
[11]    orientation_rate_z: 0.0
[12]    wind_speed: 0.0
[13]    engine_velocity_command: 0.04
[14]    rudder_angle: 0.0
[15]    Under Attack: 0.19
Next? [Y/n]
```

Attacking X & Y Position

Low Confidence Levels

Overall Threat Level

```
Under Attack: 0.19
```

Attacking X&Y Position

Policy Demonstration

- **Sanity Check**

```
Observation:
[0]     position_x: 263.44
[1]     position_y: 36.75
[2]     position_z: 0.93
[3]     orientation_x: 0.05
[4]     orientation_y: -0.01
[5]     orientation_z: 0.96
[6]     position_rate_x: 1.49
[7]     position_rate_y: 0.04
[8]     position_rate_z: 0.85
[9]     orientation_rate_x: 0.12
[10]    orientation_rate_y: -0.06
[11]    orientation_rate_z: 0.0
[12]    wind_speed: 3.0
[13]    engine_velocity_command: 200.0
[14]    rudder_angle: 0.04
Index to modify: 0
Perturbation: 123123421
Index to modify: 1
Perturbation: 98324211
Index to modify: 13
Perturbation: 0
Index to modify: no
Action:
[0]     position_x: 0.0
[1]     position_y: 0.0
[2]     position_z: 0.0
[3]     orientation_x: 0.09
[4]     orientation_y: 0.0
[5]     orientation_z: 0.0
[6]     position_rate_x: 0.0
[7]     position_rate_y: 0.01
[8]     position_rate_z: 0.0
[9]     orientation_rate_x: 0.13
[10]    orientation_rate_y: 0.02
[11]    orientation_rate_z: 0.1
[12]    wind_speed: 0.0
[13]    engine_velocity_command: 0.0
[14]    rudder_angle: 0.0
[15]    Under Attack: 0.4
Next? [Y/n]
```

Large X and Y Position

Engine Off

Low Confidence Levels

Overall Threat Level

```
Under Attack: 0.4
```

Attacking X/Y Positions and Engine Command

Policy Demonstration

- **Dashboard for easy visualization**

Roadmap

- Problem Description
- Methodology Exploration
- Reinforcement Learning (RL) Problem Formulation
- Setting up the RL Toolchain
- Policy Demonstration
- **Future Work**

Future Work

- Train policy for 10 - 50 million timesteps
- Introduce network traffic rate and time to observation space
- Develop attack modules modeled from real-world APTs
- Configure simulator to represent real-world ship

Thank you!

Developed Codebase:

https://github.com/jasonjabbour/nta_rl

Contact me:

jason.jabbour@jhuapl.edu

jasonjabbour@g.harvard.edu