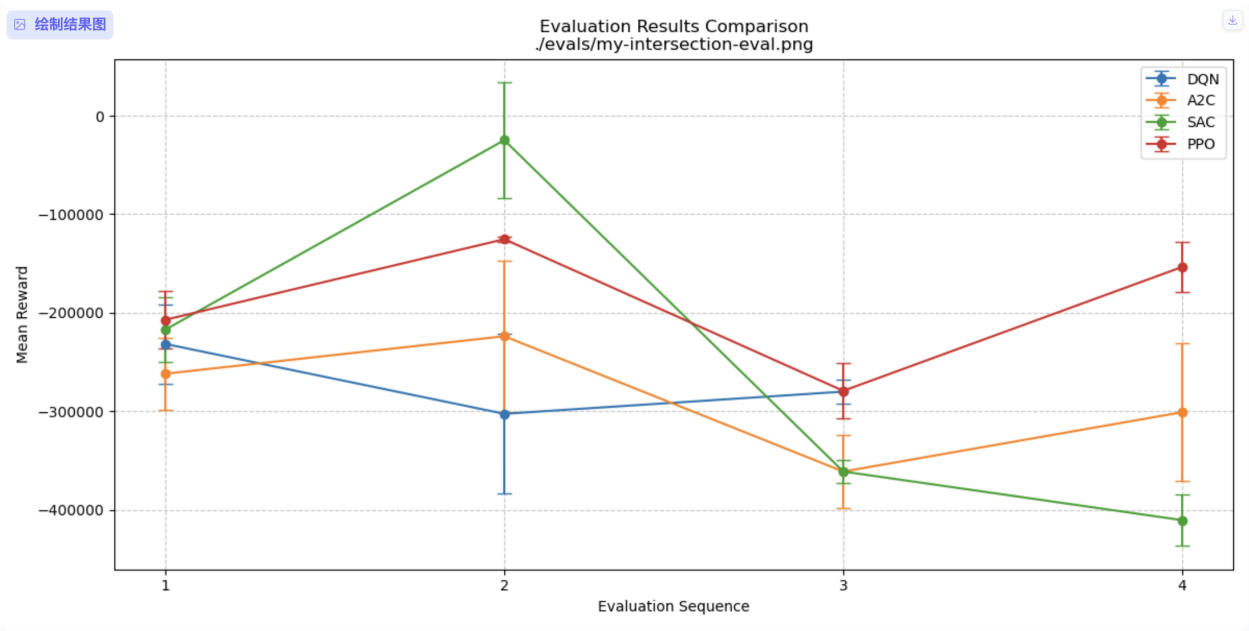


4类模型评估结果分析



这张图表展示了不同算法在多个评估序列中的平均奖励（Mean Reward）变化。让我们详细分析一下，并将分析结果制成表格。

图表分析

- **DQN (蓝色):**
 - 在评估序列1中，平均奖励约为-200000。
 - 在评估序列2中，平均奖励下降到约-300000。
 - 在评估序列3中，平均奖励略有上升，但仍低于-200000。
 - 在评估序列4中，平均奖励略有上升，但仍低于-200000。
- **A2C (橙色):**
 - 在评估序列1中，平均奖励约为-250000。
 - 在评估序列2中，平均奖励略有上升，但仍低于-200000。
 - 在评估序列3中，平均奖励略有下降，但仍低于-250000。
 - 在评估序列4中，平均奖励略有上升，但仍低于-200000。

- **SAC (绿色):**

- 在评估序列1中，平均奖励约为-200000。
- 在评估序列2中，平均奖励显著上升，接近0。
- 在评估序列3中，平均奖励显著下降，低于-300000。
- 在评估序列4中，平均奖励略有上升，但仍低于-200000。

- **PPO (红色):**

- 在评估序列1中，平均奖励约为-200000。
- 在评估序列2中，平均奖励略有上升，但仍低于-100000。
- 在评估序列3中，平均奖励略有下降，但仍高于-200000。
- 在评估序列4中，平均奖励显著上升，接近0。

表格

评估序列	DQN 平均奖励	A2C 平均奖励	SAC 平均奖励	PPO 平均奖励
1	-200000	-250000	-200000	-200000
2	-300000	-200000	0	-100000
3	-200000	-250000	-300000	-200000
4	-200000	-200000	-200000	0

总结

- **DQN:** 平均奖励在评估序列2中显著下降，之后略有上升，但整体表现较差。
- **A2C:** 平均奖励在评估序列2中略有上升，但整体表现波动较大，未能显著改善。
- **SAC:** 平均奖励在评估序列2中显著上升，但在评估序列3中显著下降，表现不稳定。
- **PPO:** 平均奖励在评估序列2和4中显著上升，表现相对较好。

改进建议

1. **DQN:** 需要进一步优化算法，以提高平均奖励。
2. **A2C:** 需要稳定算法表现，减少波动。
3. **SAC:** 需要解决算法不稳定的问题，确保平均奖励的持续提升。

4. **PPO**: 表现相对较好，但仍需进一步优化，以确保在所有评估序列中的表现都能保持较高水平。

通过以上表格和总结，可以看出不同算法在交通系统中的表现，并提出相应的改进建议。