

文献阅读报告(应用层—机器学习)

一、文献信息

1、论文题目: Increasing performance of electric vehicles in ride-hailing services using deep reinforcement learning

2、作者: Jacob F. Pettit, Ruben Glatt, Jonathan R. Donadee, Brenden K. Petersen

3、发表途径: NeurIPS 2019 Workshop Tackling Climate Change with Machine Learning

4、发表时间: 2019

二、问题意义

1、研究背景及意义

目前的交通工具所产生的二氧化碳非常多的,而且由于新型交通工具的发展可能会导致能源加速消耗,所以我们的交通工具迫切需要向无碳发电源过渡。然而无碳发电源的电动汽车却受到很多方面的限制(如行驶里程少,加油时间长,何时何地充电等问题)。这个时候就需要一个最佳的驾驶和充电策略来解决上述阻碍。

2、主要研究问题

本文探讨了利用深度强化学习(DRL)优化电动汽车代驾收费策略的可行性,旨在降低成本和排放,同时增加交通服务。作者侧重于解决当今可能采用电动汽车的独立代驾司机面临的紧迫问题,

三、思路方法

1. 建立最佳驾驶模型

作者开发了一个数据驱动的代步电动汽车驾驶模拟系统。其中包括代表终端节点控制器、电动汽车、电动汽车充电站网络和电网的模块。终端节点控制器随机生成请求的行程,行程中包括目的地、行驶距离、行驶时间和获得的运输收入。电动汽车模块跟踪电池的充电状态以便在耗尽电池或充电时重新充电,电池容量为 100kWh。电动汽车充电站网络定义了充电站的位置,并以假定的排队等待时间模式对充电站的外部电动汽车使用进行了建模。而电网根据一天中的不同时间,确定能源价格和每单位收费能源产生的二氧化碳排放量。

仿真环境基于 OpenAI Gym, 电动汽车智能体的每一步都会通过观察电动汽车电池电量、驾驶时的预期电池使用量、充电时的预期充电成本、选择充电时产生的预期排放量、车辆完成充电的预期时间,以及最近充电站的预期排队时间这几个因素,决定下一步是停下来充电还是继续驾驶。

作者用来训练智能体的强化函数为

$$r = \begin{cases} -(c + \varepsilon E) & \text{选择充电} \\ \zeta & \text{完成一次驾驶的收入} \\ -3(c + \varepsilon E) & \text{电量不足情况下坚持驾驶的惩罚} \end{cases}$$

式中, c 是为车辆充电所支付的成本, ε 是为车辆充电所产生的排放量, E 是一个用于改变智能体对车辆充电产生的排放量的考虑程度, ζ 是完成一次骑行的收入。最后这个惩罚条款可

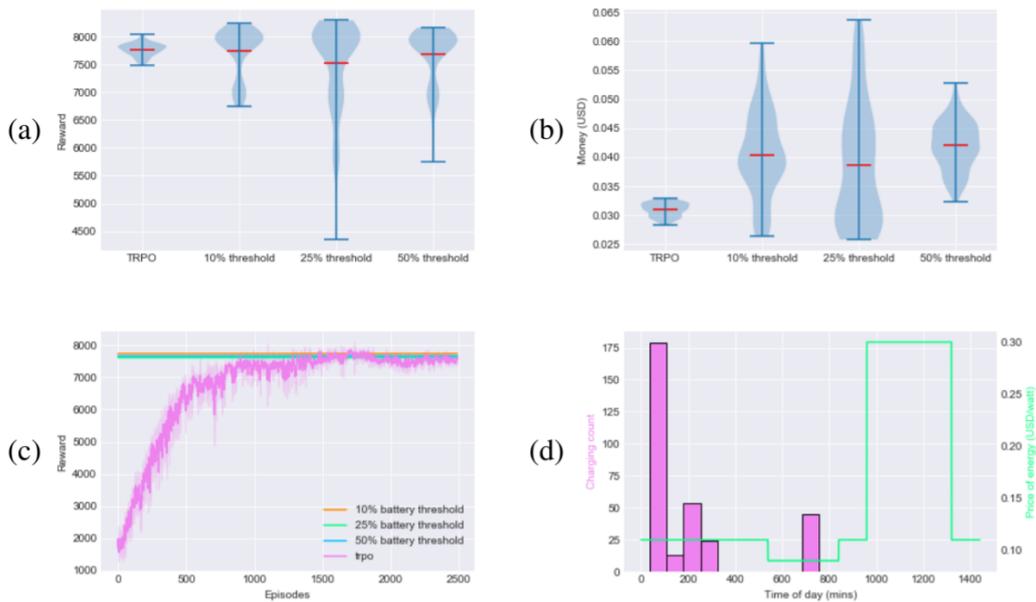
以鼓励智能体在电池耗尽前学会给车辆充电。

2. 充电策略

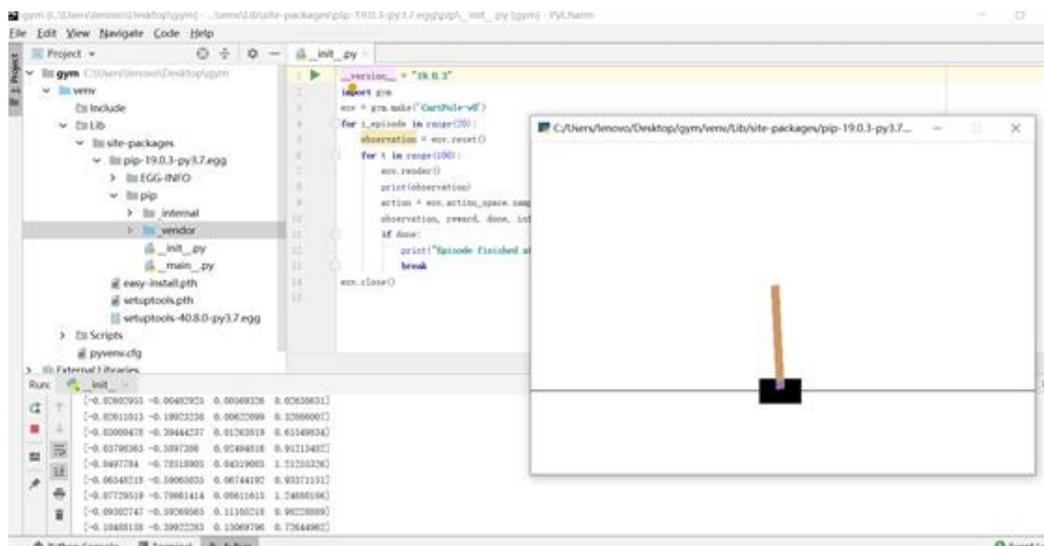
作者运用信任域策略优化 (TRPO) DRL 算法来训练适合实验环境的 EV 智能体策略, 并创建启发式策略与之对比。前者充电策略和价值函数网络是每层 64 个单位的双层感知器和双曲正切激活函数。培训批量设置为 4096, 折扣系数设置为 0.8。后者则根据当前电池电量与给定阈值的比较选择是否充电。

四、实验结论

实验结果表明经过 TRPO 算法训练的代驾电动车能够获得比启发式策略更高的平均回报, 经过训练的智能体会优先在能源价格和拥挤程度低的时候充电。



五、启发与思考



我感觉自己发现了新大陆, 这个机器学习太有意思了。虽然一开始觉得论文呈现的东西比之前的复杂的理论生动 (也有可能枯燥的部分作者没有写进论文里), 但是当我去百度了一

下文中提到的信任域策略优化 (TRPO) DRL 算法, 果然是我太天真, 不怎么能看懂。而且深度学习有一套完整的前世今生, 如果以后走这个方向得好好琢磨琢磨。可是我还是找机会感受了一下, 文中有提到仿真环境 OpenAI Gym, 手没忍住点进去一波造作, 因为比之前做别的东西的步骤简单好多。但是现在只是尝试用别人的代码感受效果, 我觉得接下来有空的话想自己试试看写代码 (虽然不太擅长), 感觉一定超有意思!