Flexibility Platform for Community Energy Systems

Shida Zhang[1], Daniel May[1], Peter Atrazhev[1], Mustafa Gul[2], Petr Musilek[1]*

[1]*Electrical and Computer Engineering, University of Alberta, Edmonton AB, Canada*
[2]*Civil and Environmental Engineering, University of Alberta, Edmonton AB, Canada*
**pmusilek@ualberta.ca*

Keywords: DISTRIBUTED ENERGY, ARTIFICIAL INTELLIGENCE, REINFORCEMENT LEARNING, TRANSACTIVE ENERGY, ENERGY FLEXIBILITY

Abstract

Integrating technological changes and sustainability considerations poses multidisciplinary challenges for the power system beyond economic and environmental benefits. Allowing energy from distributed energy resources to be traded and coordinated peer-to-peer in real-time can mitigate system and policy-making issues while decreasing the strain on power system infrastructure. TREX is an artificial intelligence (AI) assisted flexibility platform for community energy systems that can also act as an AI training tool. Using AI agents to manage instantaneous market interactions in real-time is the first step to long-term sustainability and flexibility. In this article, we show that deep learning agents are able to learn to exploit trading habits of opposing expert designed traders in a TREX environment. Based on the results, future efforts will be extended towards a multi-agent setup with full utilization of the capabilities of the market.

1 Introduction

Grid-beneficial integration of energy communities is a delicate task. On one hand, adding renewable energy (RE) and distributed energy resources (DER) without coordination could be more detrimental to the power system than their economic and environmental benefits. On the other hand, if the policies and incentives are too restrictive, then the pace of energy transition may be hindered and customers frustrated and alienated.

There is currently a lack of flexible and robust solutions to actively real-time share and coordinate DER within communities. A well-designed energy flexibility platform will mitigate both system and policy issues. Better coordination will decrease load variability and localized voltage violations, eliminating the need for roundabout symptom-oriented solutions, such as decreasing feed-in tariffs or increasing curtailment.

Our proposed energy flexibility platform, dubbed Transactive Renewable Energy Exchange (TREX), is designed to enable active, real-time energy coordination through the combination of a micro-transactive energy market, and AI-based energy management and trading agents.

The success of deploying TREX depends on the close integration of market design and agent design. To accomplish this, TREX must also be an efficient AI training environment for the participating agents. A digital twin of a real community can be constructed to realistically evaluate the effects on economics and power flow, amongst other metrics, without the need to implement slow and expensive pilot projects.

This article is organized as follows. Section 2 provides an overview of the philosophy, architecture and design of the TREX platform. Section 3 describes experiments used to demonstrate the basic abilities of proposed AI agents, while section 4 analyses their results. Major conclusions are presented in section 5. The authors assume that the reader has a basic understanding of economics, electricity markets, grid operations, machine learning (ML), reinforcement learning (RL), and AI. Background knowledge on these concepts is directly introduced, but references are provided.

2 TREX

2.1 Philosophy

A resilient, flexible, end-user focused, and DER-centric energy system must be able to effectively deal with uncertainty towards loads or feed-ins. This is a common issue with pre-optimized or scheduled systems that often results in scalability and adaptability issues. Contemporary strategies to deal with such issues are shifting responsibility towards end users through demand response (DR), as successful DR programs are highly dependent on customer education and participation [1]. Consumers are increasingly frustrated with these policies and may eventually abandon the grid. Mass defection would be very detrimental to the electricity infrastructure. To mitigate this scenario, energy coordination should be automated, real-time, and using adaptive approaches. TREX works towards achieving this goal with two complementary components:

1. Deployment of localized, real-time transactive energy markets. Given competent actors, principles of market economy [2] should naturally improve energy coordination and system balance within a distributed architecture.
2. Design and deployment of ML trained AI energy trading agents that can effectively use the aforementioned market, automating interactions for participants.

The close integration of these components should facilitate emergence of self-adaptability.

2.2 System Architecture

The TREX architecture is designed to be modular and scalable. The clients are independent processes that communicate through an socket.io server, which handles high-level tasks such as client management and message relaying. The architecture allows for operations both in deployment mode and simulation mode. To train AI agents, a digital community can be recreated with real data, allowing for a high degree of consistency.

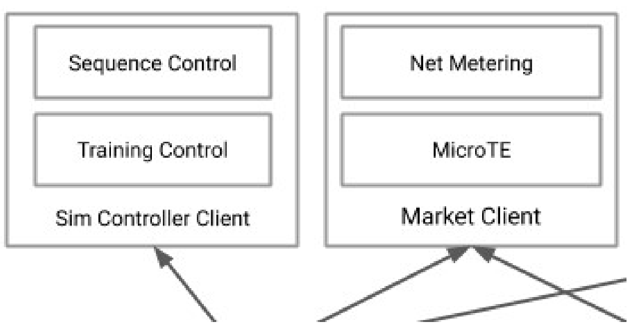

Figure 1. Simplified TREX Architecture

2.3 Market Design

Because TREX is entirely software-based, the market mechanisms must be explicit and specific. This is important as the market response for every combination of actions is deterministic.

Micro-Transactive Energy Market (MicroTE) is a market designed for TREX. Because MicroTE is small and localized, the communication overheads are minimized and can enabling higher settlement frequencies. This is especially relevant for highly DER penetrated systems with transient speeds in the order of minutes.

MicroTE assumes that the grid is always available and can be interacted with according to net-metering rules, buying for 14.49 cts/kWh and selling for 6.9 cts/kWh as per retail electricity pricing in Alberta in Nov2019. The rest of the MicroTE rules are as follows:

1. The local market has two energy pools: one for dispatchable sources (such as batteries), and one for non-dispatchable sources (such as solar).
2. Auctions settle for energy to be delivered during one-minute period from the end of the current round.
3. During the current round, participants bid and ask for energy to be delivered during or beyond the next delivery period.
4. Double auctions are used for settlements: bids/asks are settled pairwise, with bids sorted from the highest to lowest, and asks in reverse. Settlement only occurs if bid price is greater than or equal to ask.
5. Bid/ask quantities can be partially settled.
6. The settlement price is the average of bid and ask prices.
7. For hardware integration reasons, bid/ask quantity must be an integer multiple of 1 Wh to allow direct use of the watt-pulse function of most smart-meters.
8. During the delivery period, if a seller experiences a supply shortage, it must financially compensate at net metering prices. If a buyer settled for more than used, the buyer must still pay for the extra energy at settlement price.

2.4 Agent Design

The agent's behavior via market interaction is defined by setting a small set of parameters (bid/ask, price, quantity, and source) each settlement round. We chose (deep) reinforcement learning (DRL), a machine learning framework focused on self-supervised learning of optimal behaviors for sequential decision-making problems using deep neural networks. The network receives a set of observations of the environment's current state and then adjusts its behavior accordingly. The agent tries to maximize the expected cumulative reward (a feedback metric called value).

The specific algorithm used in this article is DQN [3], an established DRL algorithm. DQN's goal to learn action values for every possible action, given a set of observations. If successful, optimal acting is simply taking the action with the maximal value. Much improvements to DQN have been introduced and established. This work utilizes some that do not require the tuning of any additional parameters [4-8] to decrease complexity.

To make sure the agent experiences enough of the environment, it performs learning with an exploratory behavior strategy called ε-greedy policy, acting randomly with probability ε and otherwise following the maximal value action. During validation, the agent follows the greedy, value optimizing policy.

3 Experimental Design

With clarity and analysis in mind, we limit the scope of this work to a highly controlled environment. We solely aim to demonstrate the ability of basic deep reinforcement learning (DRL) agents to learn utilizing MicroTE, exhibit stable learning characteristics and outperform competing participants with fixed policy.

The resulting experimental setup consists of one learning agent participant and two fixed, manually designed participants. All participants are given the ability to perfectly predict generation and consumption to eliminate the influence of prediction errors. No energy storage is used and therefore only the non-dispatchable pool is relevant. The participants attempt to sell/buy energy according to residual generation and the only observations allowed for participant logic are time-of-day and day-of-week, both mapped to a unit circle.

The agent's hyperparameter choices largely correspond to values commonly reported in the literature. ADAM [9] with a batch size of 24, learning rate of 0.0001, gradient norm clipping of 20 and otherwise default parameters are used to perform optimization. The Q-network is a 3-layer fully connected architecture with 64, 128 and 256 rectified linear units (ReLU), initialized using a scheme introduced by He et. al [10].

Environment specific choices are the length of the replay buffer (56160 auction rounds), learning offset (20160 auction rounds), and network update frequency (51840 auction rounds). The agent's action space consists of the preferred bid-price and ask-price for each time-step. The output heads of the branching Q-network have linear activation and cover the action space from bid/ask-prices (0.069 to 0.1449 $/kWh) in 30 steps.

Three experiments were conducted. The goal for the first experiment is to learn the optimal bidding price only. The goal for the second and third experiments is to learn both optimal bidding and asking prices. In the first and second experiments, all non-learning agents bid and ask at fixed, time-invariant prices. The third experiment gives the non-learning agents time-variant prices, based on time-of-use. This is a reasonable representation of a simple, expert designed systems.

Experiments were performed over three stages:

1. Training was performed for one epoch (49 days of transitions).
2. Validation was performed by repeating the epoch with the currently found best strategy. The returns achieved in the validation epoch were used to evaluate agent performance.
3. To establish an upper performance ceiling, a best response for the learning agent was calculated. For all experiments this best response was static, and the determination was therefore only performed once.

4 Results and Discussion

Since the goal of this paper is to show if TREX can be used as a DRL environment, optimal convergence and convergence speed are not of major focus. Under given hyperparameters, learning appears to be stable and non-asymptotic after 100 episodes for all three experiments.

The results of the baseline experiments shown in Fig 2. and Fig. 3 clearly demonstrate that the DQN agent learns when faced with fixed price participants. The difficulty here is to learn a fixed policy that is uncorrelated with the agent's observations. The smoothed return curves show an approximately linear increase. As expected, learning bid and ask prices is a more difficult task with slower increasing expected return.

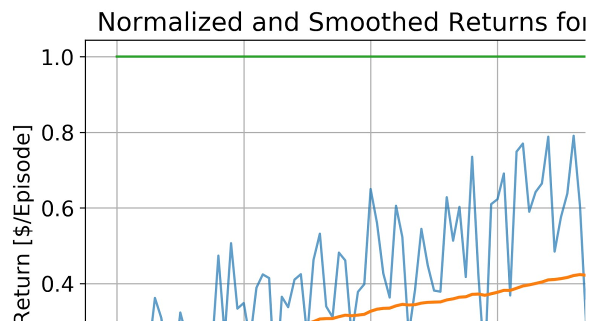

Figure 2. Normalized and Smoothed Returns for agent learning time-invariant bid price

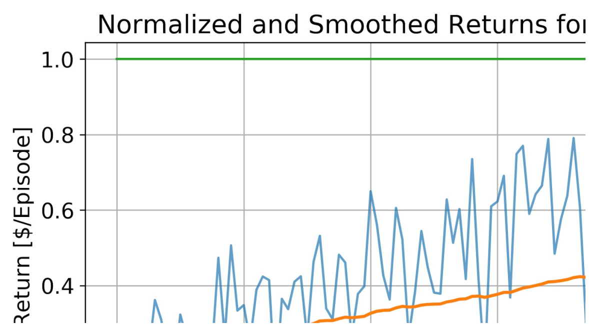

Figure 3. Normalized and Smoothed Returns for agent learning time-invariant bid and ask prices

Fig 4 shows that the DQN agent learns to outperform an expert-designed system with time-variant prices. Compared to the expert-designed system, the speed of convergence is faster. This can be attributed to a higher correlation of the return with the agent's observation space. More complicated expert-designed systems as well as DRL agents can be expected to emerge from expanded observation spaces and looser behavioral boundaries. Since more complicated scenarios such as the inclusion of batteries, imperfect forecasts and more complex observation spaces are out of the scope of this article, we leave such investigations for future work.

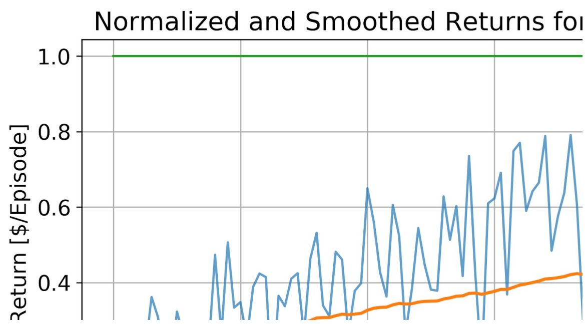

Figure 4. Normalized and Smoothed Returns for agent learning time-variant, time-of-use-based bid and ask prices

In general, as expert-designed systems are deterministic by nature, there is always an optimal response strategy that exploits part of the expert system and can be learned. For setups such as TREX, expert-designed systems are therefore a suboptimal choice. Future work will thus focus on improving DRL algorithms and training approaches for TREX specifically and partially observable, stochastic, competitive-collaborative multi-agent environments in general.

5 Conclusions

This contribution introduces TREX as a flexibility platform to improve coordination within localized energy communities that exhibit high DER penetration. TREX leverages two main components. A localized micro-transactive energy market that allows quasi real-time trading via a double-auction system, and DRL-based AI agents to automate market interaction for participants. As a realistic, model-free simulation, TREX closely mimics the information flow of a deployed system, therefore opening up the possibility to become a practical and realistic training environment for AI to automate market interactions.

TREX's viability as an AI training environment is demonstrated by running several minimalistic experiments: training a DQN agent against static and time-dependent behaviors. The agent learns in a stable manner. This clearly shows the scalability potential of the proposed flexibility platform, since the only requirement is local data.

Based on the presented results, future efforts will focus on DRL agents over expert-designed systems for market interaction, working towards a multi-agent setup with full utilization of the market's capabilities.

6 Acknowledgements

This work has been supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) grants and by the Canada First Research Excellence Fund (CFREF) under the Future Energy Systems research initiative at the University of Alberta. The authors would like to thank Dr. Andrew Leach and Dr. Tim Weis for their advice on the operation of electricity markets.

7 References

[1] Albadi, M. H., El-Saadany, E. F.: 'Demand Response in Electricity Markets: An Overview', *2007 IEEE Power Engineering Society General Meeting*, Tampa, FL, 2007, pp. 1-5.

[2] Smith, A., Cannan, E.: 'The Wealth of Nations'. (New York, N.Y., Bantam Classic, 2003, 6th edn.)

[3] Mnih, V., Kavukcuoglu, K., Silver, D. et al.: 'Human-level control through deep reinforcement learning.' Nature 518, 2015, pp. 529–533

[4] Hasselt, H. V.: 'Double Q-Learning'. Proc. Advances in Neural Information Processing Systems 23, 2010, pp. 2613–2621

[5] Zhang, S., Sutton, R. S.: 'A Deeper Look at Experience Replay' arXiv preprint, 2018, arXiv:1712.01275

[6] Tavakoli, A., Pardo, F., Kormushev, P.: 'Action Branching Architectures for Deep Reinforcement Learning' arXiv preprint, 2019, arXiv:1711.08946

[7] Wang, Z., Schaul, T., Hessel, M. et al.: 'Dueling network architectures for deep reinforcement learning.' International Conference on Machine Learning, 2016

[8] Hessel, M., Modayil, J., van Hasselt, H, et al.: 'Rainbow: Combining improvements in deep reinforcement learning'. arXiv preprint, 2017, arXiv:1710.02298

[9] Kingma, D. P., Ba, J.: 'Adam: A method for stochastic optimization.' arXiv preprint, 2017, arXiv: 1412.6980

[10] He, K., Zhang, X., Ren, S., Sun, J.: 'Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification', arXiv preprint, 2015, arXiv: 1502.01852