SCIENCE CHINA
Information Sciences

• LETTER •

November 2018, Vol. 61 110203:1–110203:3
https://doi.org/10.1007/s11432-018-9597-8

Special Focus on Optimization for Cyber-Physical Energy Systems

# On distributed event-based optimization for shared economy in cyber-physical energy systems

Qing-Shan JIA* & Junjie WU

*Center for Intelligent and Networked Systems (CFINS), Department of Automation, Beijing National Research Center for Information Science and Technology (BNRist), Tsinghua University, Beijing 100084, China*

Received 1 June 2018/Accepted 21 August 2018/Published online 18 October 2018

**Citation** Jia Q-S, Wu J J. On distributed event-based optimization for shared economy in cyber-physical energy systems. Sci China Inf Sci, 2018, 61(11): 110203, https://doi.org/10.1007/s11432-018-9597-8

Dear editor,

In order to accommodate the sharp increase in energy consumption and environmental pollution, it is important to coordinate the supply and the demand. This has become a key feature in a cyber-physical energy system (CPES), where the energy consumer and supplier share information with each other to improve the overall energy efficiency and therefore to reduce the environmental pollution. Shared economy has risen sharply across the globe and becomes an effective model to improve the efficiency in an investment. Examples include but are not limited to shared electric vehicles (EVs [1]) and bicycles. It is of great practical interest to schedule these shared objects to maximize the social welfare and sustainability.

This problem usually faces the following challenges. First, the curse of dimensionality. The state space increases exponentially with respect to (w.r.t) the system scale. Second, the curse of modeling. It is usually difficult to specify parameter settings in practical application. Third, the uncertainty. The demand and the supply in a shared economy are highly uncertain. For example, the demand on shared EVs and the availability of idle EVs depend on the individuals' behavior and are highly uncertain. Fourth, the system dynamics in multiple spatial and temporal scales. For example, there are forecasting models for renewable power generation such as wind power and solar power at different temporal scales. There also exist models to predict the demand and supply of shared EVs at different spatial resolutions. But it is not clear how to merge these models for system-level optimization.

Many efforts have been devoted to address these challenges, which can be roughly classified into two categories. First, explore structural property to reduce the state and action space. Examples include state aggregation [2] and time aggregation [3]. These methods could be very effective but are usually problem dependent. Second, develop approximate solutions to the optimal policy. Examples include approximate dynamic programming [4] and event-based optimization (EBO) [5, 6]. Among these efforts, EBO has an appealing feature since the number of events may increase only linearly w.r.t. the system scale or even stay as a constant [7].

We consider the optimization for shared economy in CPES in this study, and make the following major contributions. First, we formulate the problem as a distributed EBO to exploit the similarity among the shared objects (e.g., EVs). Second, we develop a distributed Q-learning algorithm to solve this large-scale optimization problem. Third, we demonstrate the performance of this algorithm on the scheduling of shared EVs to satisfy the trip

* Corresponding author (email: jiaqs@tsinghua.edu.cn)

 

demands integrated with renewable energies like wind power.

*Problem formulation.* To demonstrate the advantages of the proposed algorithm, we apply it to solve the problem of scheduling shared EVs introduced in [8]. It is assumed that there are several independent high-rising buildings, mounted with wind turbines that generate electricity. Wind power storage is not considered. So the generated wind power will not be stored or transferred to the grid and can only be used instantly. There is a parking lot under each building where the shared EVs are parked and charged. The batteries of the EVs can be charged either by electricity bought from the grid market or the wind power generated from the turbines. The driving demand from the users is stochastic, which is from one building to another. The goal is to maximize the income of the operator of the shared EVs by properly scheduling the EVs to pick up users and for charging. We formulate the problem as a Markov decision process (MDP) and provide more details in Appendix A.

The system state $S_t$ consists of the wind power $W_t$, the user demand $D_t$, and the state of EVs $V_t$:

$$S_t = (W_t, D_t, V_t). \tag{1}$$

The EVs' state $V_t$ consists of the state of charge (SoC) $C_t$, the ranking of SoC $R_t$, the location $L_t$, and the remaining driving/parking time $\tau_t$. The state of wind power $W_t$ is discretized to represent the number of EVs that may be charged. In order to develop a scalable problem formulation, we normalize $W_t$ and $R_t$ to [0,1] and discretize into finite (say 10) disjoint intervals.

*Event definition.* In EBO, an event $e_t$ is a set of state transition pairs:

$$e_t = \{\langle S_t, S_{t+1}\rangle\}. \tag{2}$$

We use the number of events to quantify the complexity of an event-based policy. We consider the definition of two types of events, namely the macro events $e^M$ and the micro events $e^m$, where the macro events may describe the fluctuation of the wind power and the micro events may describe the change of the ranking of the local SoC.

$$e = \{e^M, e^m\} = \{\langle W_t, W_{t+1}\rangle, \langle R_t, R_{t+1}\rangle\}. \tag{3}$$

*Distributed Q-learning for EBO.* We developed a distributed optimization framework for EBO. In this framework a central operator collects information (including wind power, user demand, the state of EVs, and reward) and distribute to the EVs. Then each EV conducts policy optimization in a distributed manner. In this study, we propose a Q-learning algorithm to solve the multi-stage decision making problem. To coordinate the global and local objectives, we introduce a weighted reward for the optimization at each EV. Detailed description on the algorithm can be found in Appendix B.

*Numerical results.* We apply the proposed method to solve a shared EVs scheduling problem with 3 buildings and 50 EVs. The experiments are run on a computer with CPU as Intel (R) Core (TM) i5-4460 and RAM memory as 8 GB. We list some important results here. More numerical results can be found in Appendix C.

We design 11 experiments with micro-events and macro-events under different complexities (i.e., discretization level of the state space). The event definitions in the 11 experiments are listed in Table C2. From Experiment Nos. 1–11, the complexity decreases. Figure 1 shows the policy performance and the average time for each iteration in each experiment.

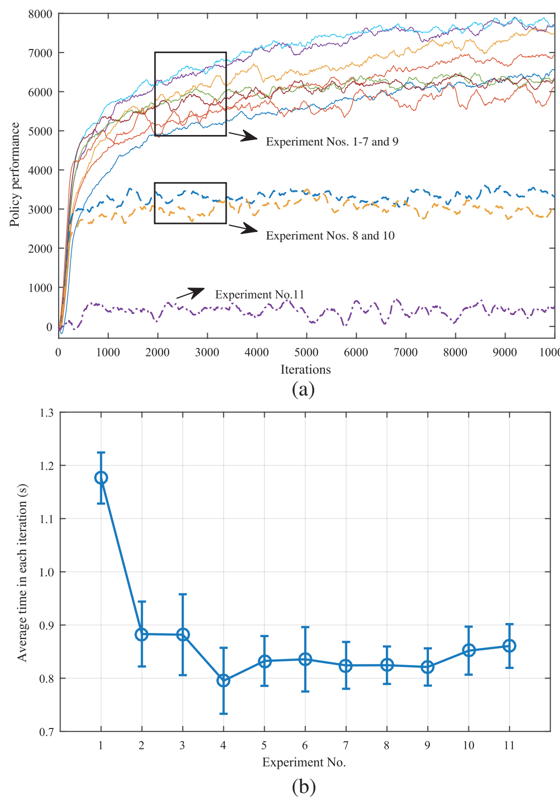


**Figure 1** (Color online) Policy performance. (a) Performance of the objective function; (b) average run time in each iteration.

The objective performances of the policy are shown in Figure 1(a) and may be classified into 3 groups. Experiment No. 11 has the poorest performance as it has the lowest event complexity. The middle are the performance curves in Experiment

Nos. 8 and 10. The micro-event $e^m$ complexity is 2 (low complexity) in both experiments. The performances are improved compared to Experiment No. 11, but are far from satisfaction. The top are a bunch of curves in Experiment Nos. 1–7 and 9. The micro-event $e^m$ complexity is larger than 5 (high complexity) in these experiments. The policy performances in this group are satisfying and is much better than the other two groups.

We further focus on the performance curves in Experiment Nos. 1–7, which can be classified into 4 groups according to the convergence rates. Experiment No. 1 with the highest event complexity has the lowest convergence rate among others. It is observed that in a certain range, it takes shorter iterations for the policy to converge to a satisfactory performance when the events become simpler.

Another important issue is the average time in each iteration, which is shown in Figure 1(b). The results suggest that the average time in each iteration may be saved by properly defining events with lower event complexity. This is because the scale of Q-table is reduced with lower complexity, which leads to faster retrieval time. While the average iteration time does not continue to decrease when the complexity drops to a certain level. This is because of the faster policy improvement rate under lower event complexity. Detailed explanation can be found in Appendix C.

Based on the numerical results shown above, we draw the remarks as follows.

**Remark 1.** Different ways to defining events may lead to large difference in policy performances.

**Remark 2.** It is possible that events with lower complexity reduce the computation budget while reserve the satisfying policy performance.

**Remark 3.** Event definitions (e.g., macro event and micro-event) even with the same complexity may have divergent policy performance.

More discussion on policy performance under various event definitions is referred in Appendix C.

*Conclusion.* We develop a distributed event-based optimization method for scheduling shared EVs in CPES. We show that the scheduling policy converges to a satisfactory performance by applying our method. Various event definitions are tested and numerical results show that the performances diverge under different event definitions. Even under the same event complexity, the way that we define events influences the performance a lot, including the policy improvement rate, the iteration time and the objective value. We show that in our shared EVs scheduling problem, micro-event achieves better performance than macro-event when the event complexity is the same. Further study may focus on event selection and quantitative analysis in distributed event-based optimization. It is also an interesting research topic to combine game theory [9] with the distributed algorithm in this study to better improve the performance of the overall system.

**Acknowledgements** This work was supported in part by National Key Research and Development Program of China (Grant No. 2016YFB0901900), National Natural Science Foundation of China (Grant Nos. 61673229, U1301254), 111 International Collaboration Project of China (Grant No. B06002), and the Program for New Star of Science and Technology in Beijing (Grant No. xx2014B056).

**Supporting information** Appendixes A–C. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

## References

1 Weiss Y, Allerhand L I, Arogeti S. Yaw stability control for a rear double-driven electric vehicle using LPV-$H_\infty$ methods. Sci China Inf Sci, 2018, 61: 070206

2 Jia Q-S. On state aggregation to approximate complex value functions in large-scale markov decision processes. IEEE Trans Automat Contr, 2011, 56: 333–344

3 Cao X R, Ren Z, Bhatnagar S, et al. A time aggregation approach to Markov decision processes. Automatica, 2002, 38: 929–943

4 Powell W B. Approximate Dynamic Programming. Hoboken: John Wiley & Sons, Inc., 2007

5 Jia Q-S, Yang Y, Xia L, et al. A tutorial on event-based optimization with application in energy Internet (in Chinese). Control Theor Appl, 2018, 35: 32–40

6 Gu Z, Huan Z, Yue D, et al. Event-triggered dynamic output feedback control for networked control systems with probabilistic nonlinearities. Inf Sci, 2018, 457: 99–112

7 Xia L, Jia Q-S, Cao X R. A tutorial on event-based optimization-a new optimization framework. Discrete Event Dyn Syst, 2014, 24: 103–132

8 Wu J, Jia Q-S. A Q-learning method for scheduling shared EVs under uncertain user demand and wind power supply. In: Proceedings of the 2nd IEEE Conference on Control Technology and Applications, Copenhagen, 2018

9 Tang C, Li X, Wang Z, et al. Cooperation and distributed optimization for the unreliable wireless game with indirect reciprocity. Sci China Inf Sci, 2017, 60: 110205