Z-learning in problems of energy systems

Aleksandr Beznosikov

Moscow Institute of Physics and Technology

Course: Machine Learning and Data Analysis
(Strijov's practice)/Group 674, 2019
Consultant: Y. V. Maximov

Problem of energy ensemble

Goal of research

We want to solve the problem in energy. For our problem there are classic methods for solving. We want to try not such known modifications and see how they will work compared to conventional methods.

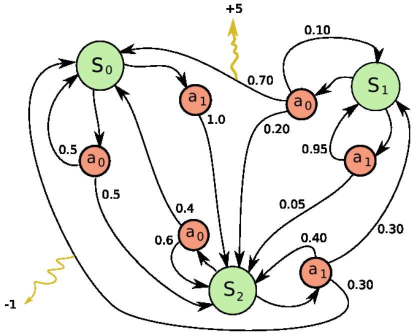
Problem

The problem occurs when the ensemble of devices for some reason turns out to be in an abnormal state. The challenge is to transfer devices to their normal state, using as less extra energy.

Method

We propose to use the classic Q-learning and its linear modification Z - learning.

$$(S, A, P, R)$$



S — state in which devices can operate

A — actions

P — all kinds of transition matrix (the desired matrix is selected depending on the action)

R — "reward" for the transition

Picture from https://ru.wikipedia.org

Main

[1] Michael Chertkov, Vladimir Y Chernyak, and Deepjyoti Deka. Ensemble control of cycling energy loads: Markov decision approach. In *Energy Markets and Responsive Grids*, pages 363–382. Springer, 2018.

[2] Csaba Szepesvári. *Algorithms for Reinforcement Learning*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, 2010.

[3] Emanuel Todorov. Linearly-solvable markov decision problems. In Bernhard Schölkopf, John C. Platt, and Thomas Hofmann, editors, *NIPS*, pages 1369–1376. MIT Press, 2006.

Auxiliary

[1] Jens Vygen Bernhard Korte. *Combinatorial optimization. Theory and Algorithms.* MCCME, translation edition, 2015.

[2] Richard Durrett. *Probability: theory and examples.* Duxbury Press, Belmont, CA, second edition, 1996.

[3] Chi Jin, Zeyuan Allen-Zhu, Sébastien Bubeck, and Michael I. Jordan. Is q-learning provably efficient? *CoRR*, abs/1807.03765, 2018.

[4] Sunyong Kim and Hyuk Lim. Reinforcement learning based energy management algorithm for smart energy buildings. *Energies*, 11:2010, 08 2018.

[5] Elizaveta Kuznetsova, Yan-Fu Li, Carlos Ruiz, Enrico Zio, Keith Bell, and Graham Ault. Reinforcement learning for microgrid energy management. *Energy*, 59:133–146, 05 2013.

[6] Raju Leo, R S Milton, and S Sibi. Reinforcement learning for optimal energy management of a solar microgrid. pages 183–188, 09 2014.

$$\min_{p,\rho} \sum_{t=0}^{T-1} \sum_{j} \rho_j(t) \left(\sum_i p_{ij}(t) \left(U_i(t+1) + \gamma_{ij} \log \frac{p_{ij}}{\bar{p}_{ij}} \right) \right) \tag{1}$$

$$\text{s.t.} \ \sum_i p_{ij}(t) = 1, \quad \forall t, \ \ \forall j \tag{2}$$

$$\rho_i(t+1) = \sum_j p_{ij}(t)\rho_j(t) = 1, \quad \forall t, \ \ \forall i \tag{3}$$

$$\text{initial conditions: } \rho_i(0), \quad \forall i \tag{4}$$

Basic code

Algorithm 1 Z - learning

Initialize: Z
Input: $\mathfrak{G} = (i_k, j_k, u_k)$
for $t = 0, 1, 2, \ldots, k-1$
    $Z(s_k) := (1 - \alpha_t) Z(i_k) + \alpha_k \exp(-u_k) Z(j_k)$
return Z

Algorithm 3 Q - learning

Initialize: Q
Input: $\mathfrak{G} = (i_k, j_k, u_k)$
for $t = 0, 1, 2, \ldots, k-1$
    $Q(i_k, p_k) := (1 - \alpha_k) Q(i_k, p_k) + \alpha_k \min\limits_{p'} (l_k + Q(j_k, p'))$
return Q

Results