Towards autonomous and optimal excavation of shield machine: a deep reinforcement learning-based approach

Ya-kun ZHANG¹, Guo-fang GONG¹**, Hua-yong YANG¹, Yu-xi CHEN¹, Geng-lin CHEN²

¹State Key Laboratory of Fluid Power and Mechatronic Systems, Zhejiang University, Hangzhou 310027, China
²School of Electrical and Power Engineering, China University of Mining and Technology, Xuzhou 221116, China

**E-mail: gfgong@zju.edu.cn

The pseudo-code for the implementation of the training environment

Input: Action {, } 
Output: Observation {, , , , , , , reward } {True or False}

* Initialization: 
  1. Load the geological data look-up table; 
  2. Load the machine-ground interaction DNN model; 
  3. Step number \( k \leftarrow 0 \); 
  4. Output the initial observation to the DRL agent; 

* function step (, ) : 
  1. \( T_i \leftarrow T_i \), \( F_i \leftarrow F_i \); 
  2. if \( k < \) the number of training steps: 
    step number \( k \leftarrow k+1 \); 
  else: 
    \( k \leftarrow 0 \); 
  end if 
  3. Read geological data \( \{P_{gw,i}, c_k, \phi_k\} \); 
  4. Scale the actual action \( \{T_i, F_i\} \) and geological data \( \{P_{gw,i}, c_k, \phi_k\} \) by dividing by their corresponding maximum values; 
  5. Calculate \( \{n_{i,j}, \dot{x}_i\} \) using the machine-ground interaction DNN model; 
  6. Scale \( \{n_{i,j}, \dot{x}_i\} \) by dividing by their corresponding maximum values; 
  7. Calculate the reward \( r_k \) value using Eq. (31); 
  8. Done \( \leftarrow \) is done (, , ); 
  return observation, reward, Done 
end function