Plant and Soil manuscript No.
(will be inserted by the editor)

Semi-active control of a structure using magneto-rheological dampers based on the reinforcement learning

Yang Yang[1]

Received: date / Accepted: date

Abstract **Keywords** Herb

1 Introduction

2 The model of motion of structures

The equation of motion of a controlled building with n degree of freedom subjected to an earthquake load can be expressed as:

$$M\ddot{X}(t) + C\dot{X}(t) + KX(t) = -M\Gamma\ddot{x}_g + B_s F(t) \tag{1}$$

where M, C, and K are the $n \times n$ mass, damping, and stiffness matrices of a controlled building. $X(t)$, $\dot{X}(t)$, and $\ddot{X}(t)$ the displacement, velocity, and acceleration with respect to the base. F(t) is the force produced by the damper. $\ddot{x}_g$ is the acceleration of an earthquake.

For Equation2, the corresponding state space equations can be expressed as:

$$\dot{x} = Ax + Bu \tag{2}$$

$$y = Cx + Du \tag{3}$$

where

$$A = \begin{bmatrix} 0 & I \\ -M^{-1}k_s & -M^{-1}C \end{bmatrix} \tag{4a}$$

$$B = \begin{bmatrix} 0 & 0 \\ -1 & -M^{-1}\Gamma \end{bmatrix} \tag{4b}$$

∗ Yang Yang
E-mail: smiba@qq.com

1 College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China

$$C = \begin{bmatrix} I & 0 \\ 0 & I \\ -M^{-1}K_s & -M^{-1}C_d \end{bmatrix} \tag{4c}$$

$$D = \begin{bmatrix} I & 0 \\ 0 & I \\ -M^{-1}K_s & -M^{-1}C_d \end{bmatrix} \tag{4d}$$

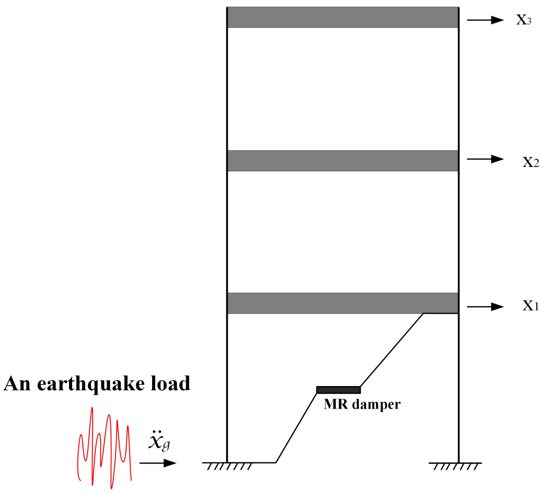

Fig. 1 Three-story building with an semi-active control system.

3 The model of MR damper

The schematic model is shown in Figure 2

The Bouc-Wen model is written as

$$F = c_0\dot{x} + k_0(x - x_0) + \alpha z \tag{5}$$

where α is the Bouc-Wen model parameter related to the MR material yield stress; k_0 and c_0 are spring stiffness and dashpot damping coefficient respectively. $\dot{z}$ and z are hysteretic deformation of the model which is defined by following equation:

$$\dot{z} = -\gamma|\dot{x}|z|z|^{n-1} - \beta\dot{x}|z|^n + A\dot{x} \tag{6}$$

where A, β, and γ are the Bouc-Wen model parameters.

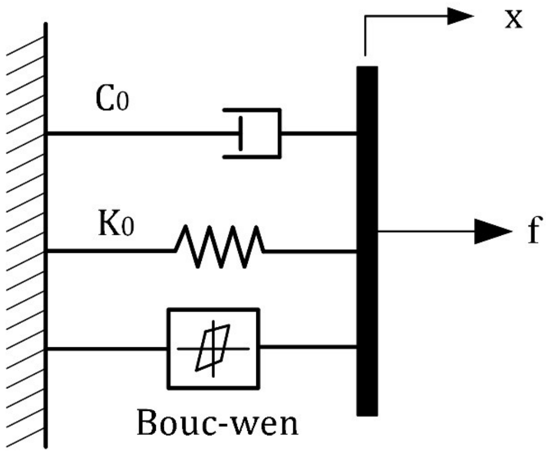

Fig. 2 The schematic diagram of the Bouc-Wen model

4 Reinforcement learning control method

Reinforcement learning is a type of machine learing that can develop solutions to sequential decision-making problems. Controlling MR dampers in the semi-active control can be modeled as a sequential decision-making proble. RL learns how to act in an environment to maximize a numerical reward signal, representing success at achieving a desired goal.

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha[R_{t+1} + \gamma max_a Q(S_{t+1,a}) - Q(S_t, A_t)] \tag{7}$$

where $\alpha \in$[0,1] is the learning rate, the ; $\gamma \in$[0,1] is the discount rate. The discount rate determines the present value of future rewards. A_t is the action set. S_t is the state set.

4.1 Action

In this study, we control the current of a MR damper to change its damping force. Hence, we selected the inputing the current for the MR damper as the action of the agent A_i.

4.2 Reward

In this study, the reward was defined in

$$R(U_i) = \begin{cases} -C\frac{U_{coni}}{U_{unconi}}, & \frac{U_{coni}}{U_{unconi}} > 1 \\ 0, & \frac{U_{coni}}{U_{unconi}} = 1 \\ C\frac{U_{unconi}}{U_{coni}}, & \frac{U_{coni}}{U_{unconi}} < 1 \end{cases} \tag{8}$$

where C is a magnification factor of the reward; U_{coni} is the structure response of controlled; U_{unconi} is the structure response of uncontrolled. When $\frac{U_{coni}}{U_{unconi}}$, the action had not change the structure response. The action had not reduece the structure response, hence, its reward was zero. When $\frac{U_{coni}}{U_{unconi}} > 1$, the action increased the structure response. Hence, its reward (punishment) is $-C\frac{U_{coni}}{U_{unconi}}$. When $\frac{U_{coni}}{U_{unconi}} < 1$, the action decreased the structure response. Hence, its reward is $C\frac{U_{unconi}}{U_{coni}}$. Thus, if an action can reduce more response, it can achieve more reward. In contrary, if an action increased more response, it will obtain more punishment. RL can give its reward according to each action using the reward function. Finally, the optimal action of each step is determined to form the optimal control strategy and realize the optimal control of MR damper.

4.3 Greedy method

$$\pi(a|s) = \begin{cases} 1 - \varepsilon + \frac{\varepsilon}{|A(s)|}, a = \arg\max_a Q(s,a) \\ \frac{\varepsilon}{|A(s)|}, \quad a \neq \arg\max_a Q(s,a) \end{cases} \tag{9}$$

5 Numerical example

The performance of the proposed reinforcement learning control algorithm was evaluted in one example through numerical simulation. To demonstrate the propsed control algorithm, a three-story building with a single MR damper was considered. The MR damper was installed in the first floor, as shown in 1. The mass M, dampering Cs and stiffness Ks matrices are as follows:

$$M = \begin{bmatrix} 3.456 & 0 & 0 \\ 0 & 3.456 & 0 \\ 0 & 0 & 3.456 \end{bmatrix} \times 10^5 kg \tag{10}$$

$$Cs = \begin{bmatrix} 1.745 & -0.512 & -0.111 \\ -0.512 & 1.634 & -0.623 \\ -0.111 & -0.623 & 1.122 \end{bmatrix} \times 10^5 Ns/m \tag{11}$$

$$Ks = \begin{bmatrix} 2.4 & -1.2 & 0 \\ -1.2 & 2.4 & -1.2 \\ 0 & -1.2 & 1.2 \end{bmatrix} \times 10^8 N/m \tag{12}$$

6 Results and Discussions

6.1 The effect of reward function

6.2 The effect of RL parameters

6.3 The effect of semi-active control by RL

7 Conclusions

References

Arnone E, Caracciolo D, Noto LV, Preti F, Bras RL (2016) Modeling the hydrological and mechanical effect of roots on shallow landslides. Water Resources Research 52(11)

Cho SE, Lee SR (2002) Evaluation of surficial stability for homogeneous slopes considering rainfall characteristics. Journal of Geotechnical and Geoenvironmental Engineering 128(9):756–763

Genuchten MTV (1980) A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Science Society of America Journal 44(44):892–898

Lu N, Godt J (2008) Infinite slope stability under steady unsaturated seepage conditions. Water Resources Research 44(11):2276–2283

Lu N, Godt JW (2013) Hillslope hydrology and stability. Cambridge University Press

Osman N, Barakbah S (2006) Parameters to predict slope stabilitysoil water and root profiles. Ecological Engineering 28(1):90–95

Pedersen A, Zhang K, Thorup-Kristensen K, Jensen LS (2009) Modelling diverse root density dynamics and deep nitrogen uptakea simple approach. Plant & Soil 326(1-2):493–510

Pollen N (2007) Temporal and spatial variability in root reinforcement of streambanks: accounting for soil shear strength and moisture. Catena 69(3):197–205

Romano N, Palladino M, Chirico GB (2011) Parameterization of a bucket model for soil-vegetation-atmosphere modeling under seasonal climatic regimes. Hydrology and Earth System Sciences 15(3):3877–3893

Wu J, Zhang R, Gui S (1999) Modeling soil water movement with water uptake by roots. Plant & Soil 215(1):7–17

Zhang HR, Fa-Hu LI, Zhang XP (2015) Effects of wild herbaceous vegetation roots on undisturbed surface soil shear strength. Journal of China Agricultural University 20(4):189–195