EPiC Series in Engineering
Volume 3, 2018, Pages 1465–1469
HIC 2018. 13th International Conference on Hydroinformatics

Autonomous Control of Urban Storm Water Networks Using Reinforcement Learning

Abhiram Mullapudi[1], Branko Kerkez[1]

[1] Department of Civil Engineering, University of Michigan, Ann Arbor, Michigan-48105, United States

Corresponding author: bkerkez@umich.edu

Abstract. We investigate the real-time and autonomous operation of a 12 km^2 urban storm water network, which has been retrofitted with sensors and control valves. Specifically, we evaluate *reinforcement learning,* a technique rooted in *deep learning,* as a system-level control methodology. The controller opens and closes valves in the system, which enhances the performance in the storm water network by coordinating the discharges amongst spatially distributed storm water assets (i.e. detention basins and wetlands). A reinforcement learning control algorithm is implemented to control the storm water network across an urban watershed. Results show that control of valves using reinforcement learning shows great potential, but extensive research still needs to be conducted to develop a fundamental understanding of control robustness. We specifically discuss the role and importance of the reward function (i.e. heuristic control objective), which guides the autonomous controller towards achieving the desired water shed scale response.

Keywords: Real-Time Control, Reinforcement-Learning, Storm-Water Networks

1 "Smarter" Storm Water Networks

"Smarter" storm water systems, equipped with sensors and actuators, have the potential to augment static infrastructure and control its response in real time to achieve watershed-scale performance benefits [3]. It has been established, however, that control actions at the scale of individual sites (e.g. storing or discharging water at a pond, for example) do not necessarily guarantee an improvement at the system scale [4]. Any individual control action has the potential to damage or flood the system if not carried out in the context of the broader urban watershed. A watershed-level

control approach is thus needed to coordinate the flows across the scale of entire storm water system. While a number of real-time control approaches have been evaluated in the past, our new approach evaluates recent advances in *reinforcement* and *deep learning* to control storm water systems.

2 Reinforcement Learning

Reinforcement learning is a *closed loop* control approach, adopted from the artificial intelligence community. Unlike *supervised learning*, this approach to control does not force the learner (i.e. agent) to choose an action, rather it directs the agent based on heuristics (i.e. reward functions), towards discovering the optimal choice of action [2].

During a storm event, a controller trained using deep reinforcement learning will observe the state of the spatially distributed storm water assets (e.g. water levels or flows in pipes and basins) and implements a control action to drive the system towards a desired state. To this end, our approach uses Deep-Q-Learning [1] to learn the state-action mapping $Q\ (state, action)$. The approach relies on deep neural networks ($parmater\ \theta$) to approximate the action value function using Eq. [1].

$$Q(s_t, a|\theta) = Q(s, a|\theta) + \gamma\big(r + max_a Q(s_{t+1}, a|\theta) - Q(s, a|\theta)\big) \quad (1)$$

Reward, represented by "r" in Eq.1, acts as the heuristic that guides the learning process. The value of reward received by the controller during the training process is dependent on the state of the system at each time step. Formulating the reward function ($R(s)$) is one of the crucial aspects of reinforcement learning. It can not only effect the speed of learning process, but also influence the trajectory of the optimal control actions.

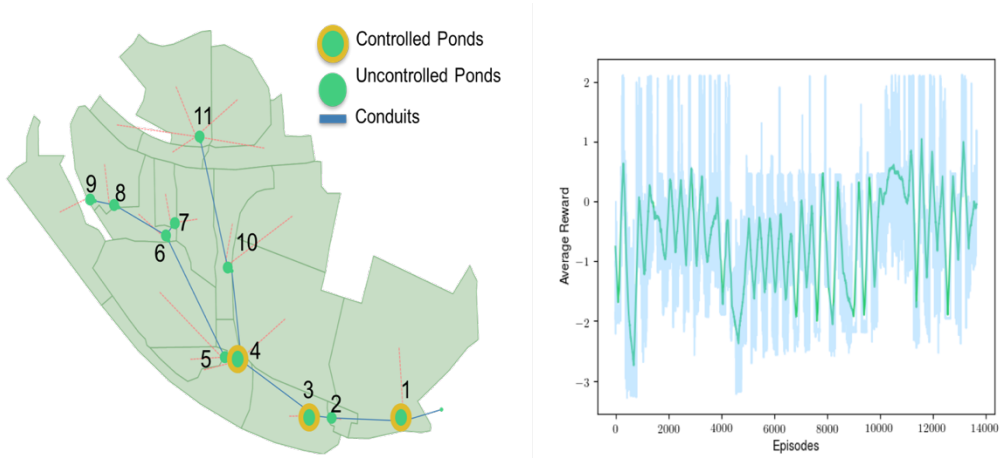

Figure.1:(a) 5sq.mile storm water network retrofitted for real time control. (b) Plot illustrating the training of the controller over 14000 episodes, where controller learns to maximize the average reward per episode.

3 System-level Storm Water Control

Here, we present the simulation-based analysis of a real-world inspired storm water network controlled in using reinforcement learning. Basins 1,3,4 (Fig.1) are assumed to be retrofitted with valves that can be controlled in real time, while rest of the system is uncontrolled. It is assumed that the controlled has access to all water levels and sensors in the system. The control objectives are formulated to discharge water from the watershed as fast as possible, without causing flooding or exceeding a critical outflow level $Q_{Critical}$. The latter constraint seeks to maximize stream stability and erosion. These objectives are embedded into a Reward Function, which guides the learner during the neural network training procedure: $\alpha, \beta, \gamma, \Gamma$ are the scaling parameters

$$R(water\ levels) = [1.0\ 1.0\ 1.0\ 1.0] \cdot [h_1\ \ h_2\ \ h_3\ h_4]^T \tag{2}$$

$$h_i = -\alpha \times height_i\ if\ hegh t_i \leq H_{critical}\ else - \gamma \times height_i^2 + \beta \tag{3}$$

$$R(flooding) = [1.0\ 1.0\ 1.0\ 1.0] \cdot [f_1\ \ f_2\ \ f_3\ f_4]^T \tag{4}$$

$$f_i = 10.0\ if\ flooding_i > 0.0 \tag{5}$$

$$R(outflows) = [1.0\ 1.0\ 1.0\ 1.0] \cdot [o_1\ \ o_2\ \ o_3\ o_4]^T \tag{6}$$

$$o_i = 1.0\ if\ outflow < Q_{critical}\ else\ \ (Q_{critical} - outflow_i) \times \Gamma \tag{7}$$

$$R(t) = R(water\ level) + R(flooding) + R(outflows) \quad (8)$$

The controller was trained over 14000 instances of storms on NVIDIA Tesla K20 GPUs, to learn the state-action value function of the watershed based on a number of design storms (25 year events). Figure.1(b) illustrates the growth in the average reward earned per episode during the training process. During each time step, the state-action value function observes the water levels in the basins (Fig.1) and previous valve positions to determine the action-value. The action in encoded as a scalar value for a valve-positions that drives the network towards desired behaviour. The controller picks the action associated with the maximum predicted action-value.

4 Results

The reinforcement learning-based controller largely maintained the flow in the network below the critical threshold without causing flooding. As such, it shows great promise as a potential control tool for storm water systems. However, a number of instances and behaviours warrant further investigation. For example, the outflow from the basin 1 exceeded the maximum threshold by almost a factor of 2. This deviation was a result of the choice of the weights in the reward function [Eq.6], as all the outflows from the ponds were equally weighted. The controller received a positive reward for maintaining the flows in the ponds 2,3,4 below the maximum flows. This positive reward overpowered the negative reward caused by the deviations in pond 1. A stricter bound on the outflows in the downstream pond could be enforced by altering the weights (e.g. increasing the weight from 1.0 to 2.0) in the reward function. This result illustrates the importance of reward function shaping in the learning process, since the reinforcement learning agent can take advantage of potential loop holes in the reward function formulation to maximize the reward it gains. Robustness against such a behaviour should be investigated in the future.

The system-level controller was able maintain the flows in the system in a desired region by just controlling an upstream pond 4 (Fig.2). While it did throttle the valves on basins 1 and 3, the benefits were marginal. While reward function parametrization could have influenced this behaviour, it also suggests that reinforcement learning does not need to control the entire system to achieve system-level benefits.

5 Conclusion/Future Work

Our initial case study sought to frame the performance boundaries of a controlled storm water network and identify the critical elements that limit the performance. Future work is needed to determine if the behaviours observed herein could have been driven by the specific intensity, duration or distribution of the storms used in this study.

As such, while reinforcement learning shows promise for the real-time control method for water networks, its stability and generalizability have to be analysed further.

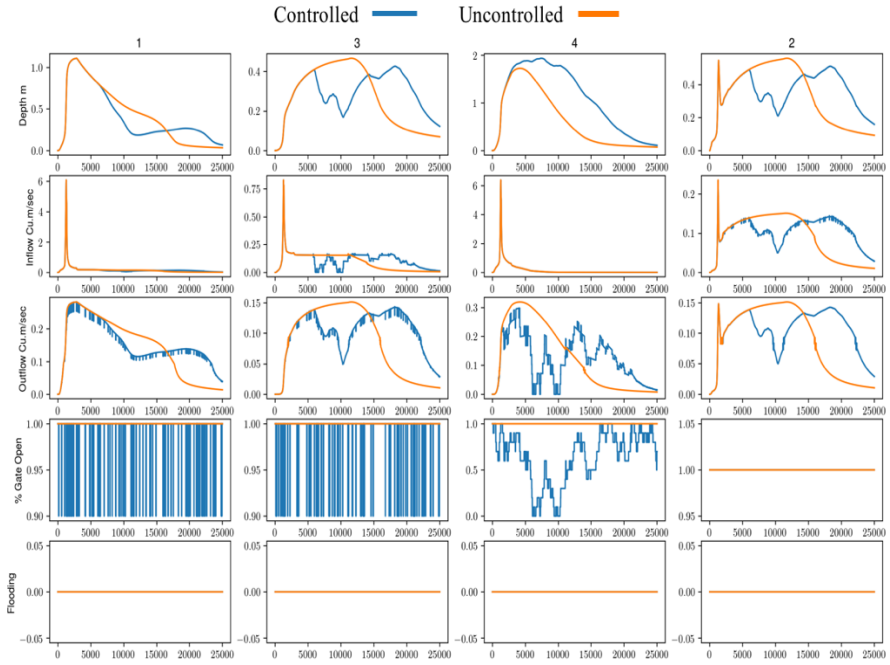

Figure 2: Plot comparing the controlled and uncontrolled response of 4 basins in the watershed during the storm event.

References

[1] Mnih, V. K. (2015). Human-level control through deep reinforcement learning. *Nature*(7540), 529-533.

[2] Sutton, R. S. (1998). *Reinforcement learning: An introduction* (Vol. 1). MIT press Cambridge.

[3] Kerkez, B. a. (2016). Smarter stormwater systems. *Environmental Science and Technology*, 7267-7273

[4] Mullapudi, A. a. (2017). Emerging investigators series: building a theory for smart stormwater systems. *Environmental Science: Water Research \& Technology, 3*(1), 66-77.