Deep reinforcement learning for rehabilitation planning of water pipes network

Zaharah A. Bukhsh, Nils Jansen, Hajo Molegraaf

Background

- Reliable water distribution systems (WDS) are paramount for functioning societies.

- Different kinds of maintenance strategies, such as recurring schedules (planned) or run-to-failure (corrective), are implemented to keep these assets from failing.

- Recurring schedules run into problem of undue maintenance, thus resulting in loss of functional life of an asset.

- Corrective solutions are expensive due to immediate need of resources and impact on surroundings (including traffic and households).

Motivation

- Typical optimization solutions are site-specific and do not include comprehensive criteria (such as economic, social impact) of large-scale pipe network.

- We propose a scalable solution to learn optimal rehabilitation policy for continuously deteriorating water pipes under economic and reliability requirements.

- We utilize deep reinforcement learning framework for development of rehabilitation policy for water pipes.

Approach
Problem formulation

- The goal is to find optimal intervention moments such that we incur the minimum average economic cost and maximum average reliability of water pipes within a planning time period.

- We model the rehabilitation planning problem of water pipes as a finite Markov decision process (MDP) within the DRL framework.

- Custom environment formulation

Approach

Custom environment formulation

- State
- Actions
- Reward
- Dynamics (time to transition)

Approach

Custom environment formulation

- State
- Actions
- Reward
- Dynamics (time to transition)

State

- Age
- Material
- Length
- Failure probability
- Reliability level
- Failure rate with respect to material and length (Wols, et. al, 2019)
- Auxiliary age (aux_t)

Failure probability of a component is its **reliability**, expressed as an exponential (Poisson) distribution

$$R = e^{-ut}$$

Where u is average failure rate

Then the probability of failure (unreliability) is

$$\mathrm{P(t)} = 1 - R(t) = 1 - e^{-ut}$$

Approach

Custom environment formulation

- State
- Actions
- Reward
- Dynamics (time to transition)

Action

- Do Nothing
- Maintain
- Replace

Reward

$$R(s_t, a_t, s_{t+1}) = \begin{cases} 0 & \text{if } a_i^t = \text{Do nothing} \\ -0.8 & \text{if } a_i^t = \text{Replace} \\ -0.5 & \text{if } a_i^t = \text{Maintain and } pf > 0.5 \\ -1 & \text{if } a_i^t = \text{Maintain and } pf \leq 0.5 \\ -1 & \text{if } a_i^t = \text{Do nothing and } pf \geq 0.9 \end{cases}$$

Approach

Custom environment formulation

- State
- Actions
- Reward
- Dynamics (time to transition)

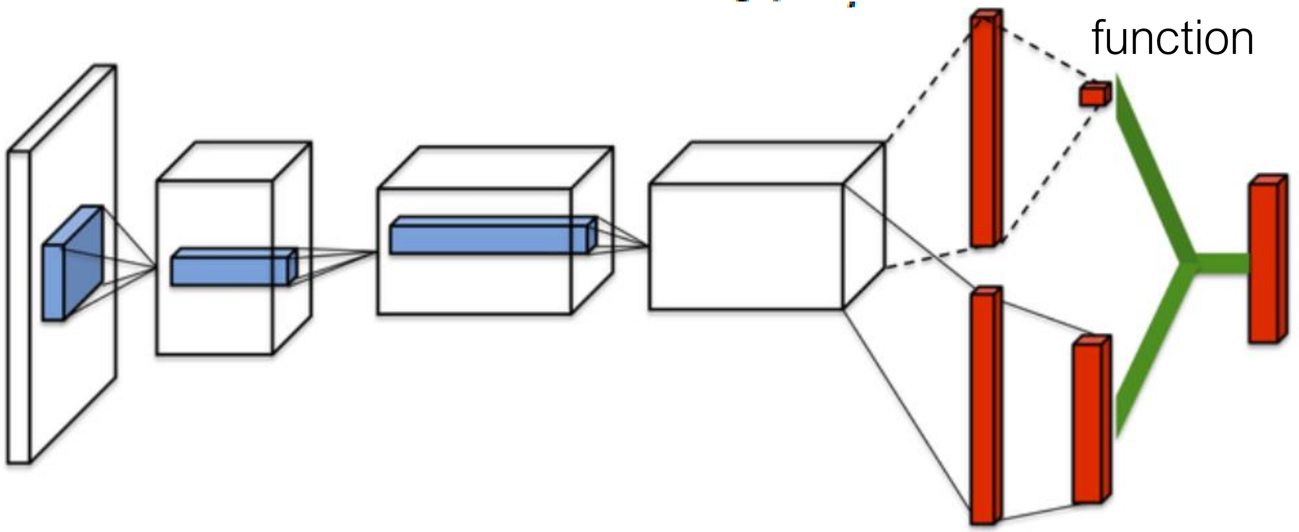
At any timestep, the agent receives a representation of the environment's state in the form of state s_t ∈ S where s_t = ⟨age,mat, L, FR, aux, pf , rl⟩. The age of the pipe is increased by one year and the aux_t is updated depending on the chosen actions as follows:

$$aux_{t+1} = \begin{cases} aux_t + 1 & \text{if } a_i^t = \text{Do nothing} \\ aux_t - 10 & \text{if } a_i^t = \text{Maintain} \\ aux_t = 1 & \text{if } a_i^t = \text{Replace} \end{cases}$$

Approach
Learning Agent

Dueling Deep Q Network

$$Q(s,a) = V(s) + \left(A(s,a) - \frac{1}{\mathcal{A}} \sum_{a} A(s,a)\right)$$



Ziyu Wang, Tom Schaul, Matteo Hessel, Hado Hasselt, Marc Lanctot, and Nando Freitas. 2016. Dueling network architectures for deep reinforcement learning. In International conference on machine learning. PMLR, 1995–2003.

Case study

“To develop and implement a DRL framework to automatically devise an optimal rehabilitation policy for water pipes network under economic and reliability requirements.

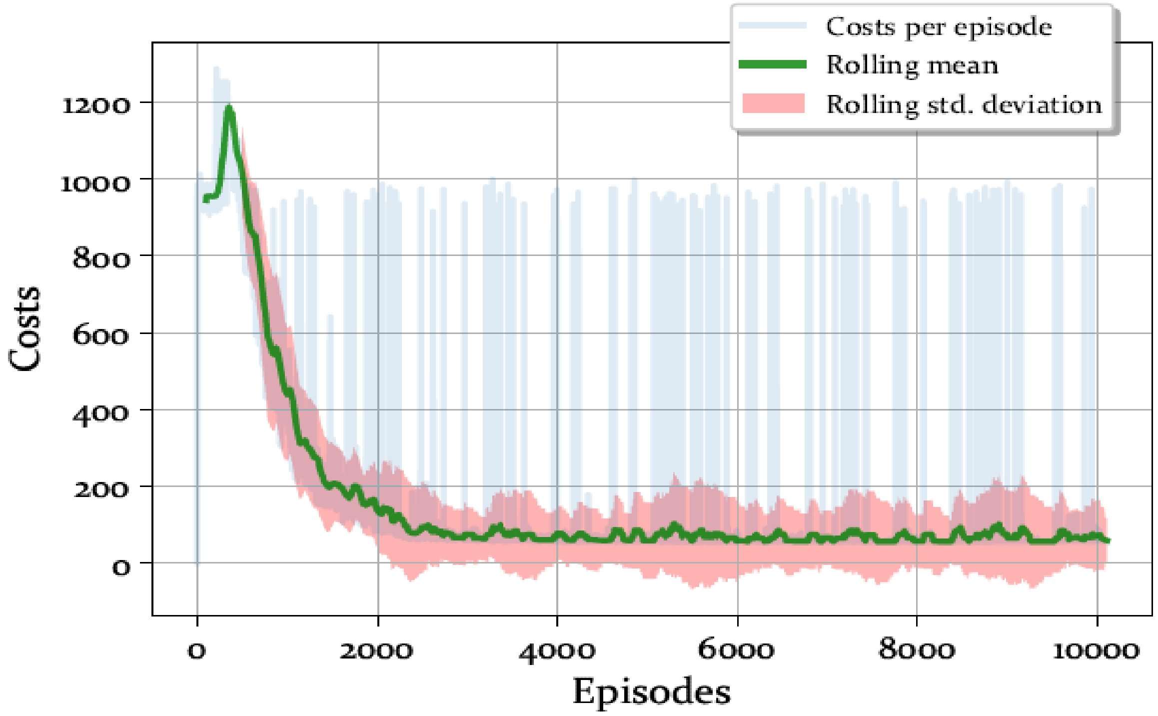
Pipe ids	Age (year)	Material	Length (m)	*Failure rate (/km)
P1	44	MAT_AC	2365	0.06
P2	46	MAT_AC	2732	0.06
P3	6	MAT_AC	1908	0.06
P4	42	MAT_AC	1996	0.06
P5	32	MAT_DI	1968	0.09
P6	37	MAT_DI	2915	0.02
P7	25	MAT_DI	2405	0.02
P8	47	MAT_DI	1500	0.02
P9	11	MAT_GCI	2017	0.07
P10	30	MAT_GCI	1679	0.07
P11	31	MAT_GCI	2071	0.07
P12	45	MAT_GCI	1666	0.07
P13	15	MAT_PVC	1650	0.015
P14	40	MAT_PVC	2365	0.015
P15	22	MAT_PVC	2434	0.015
P16	2	MAT_PVC	1527	0.015

* Wols, B. A., Vogelaar, A., Moerman, A., & Raterman, B. (2019). Effects of weather conditions on drinking water distribution pipe failures in the Netherlands. *Water Supply*, *19*(2), 404-416

Case Study

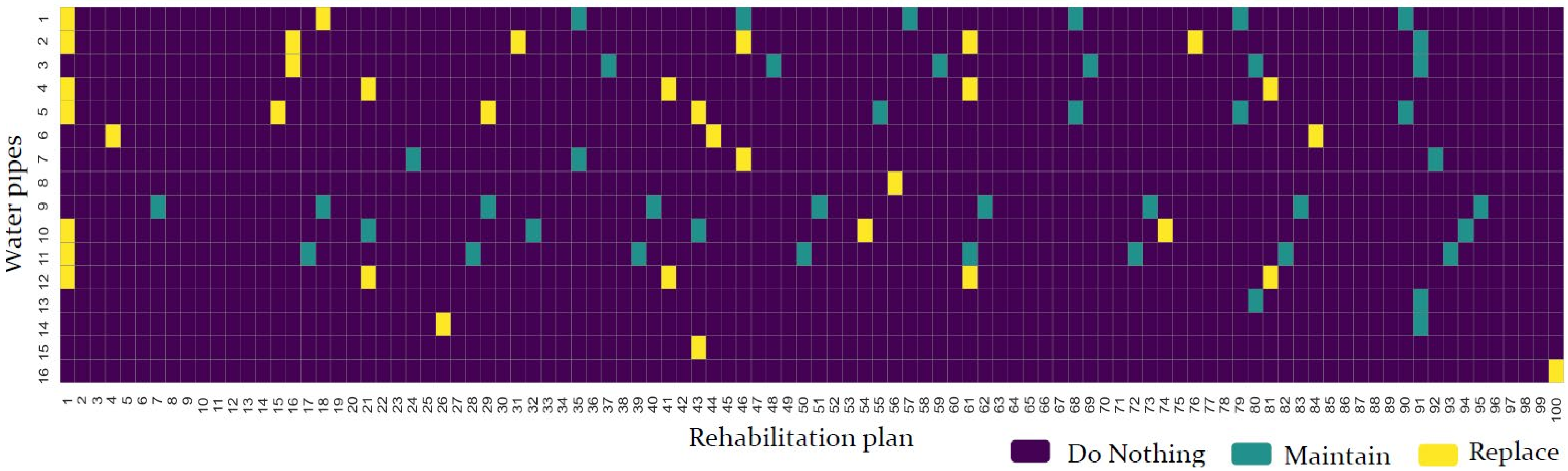
Results

Learning performance of the DDQN agent during training. The agent shows the high cost in the beginning due to randomly exploring the environment. After around 2000 episodes, the agent shows steady performance.



The learning performance of DQN agent (during training) for 10000 episodes with 100 timesteps (years). The performance stabilizes after 2000 episodes. (Smaller cost is preferred).

Deep RL approach
Results

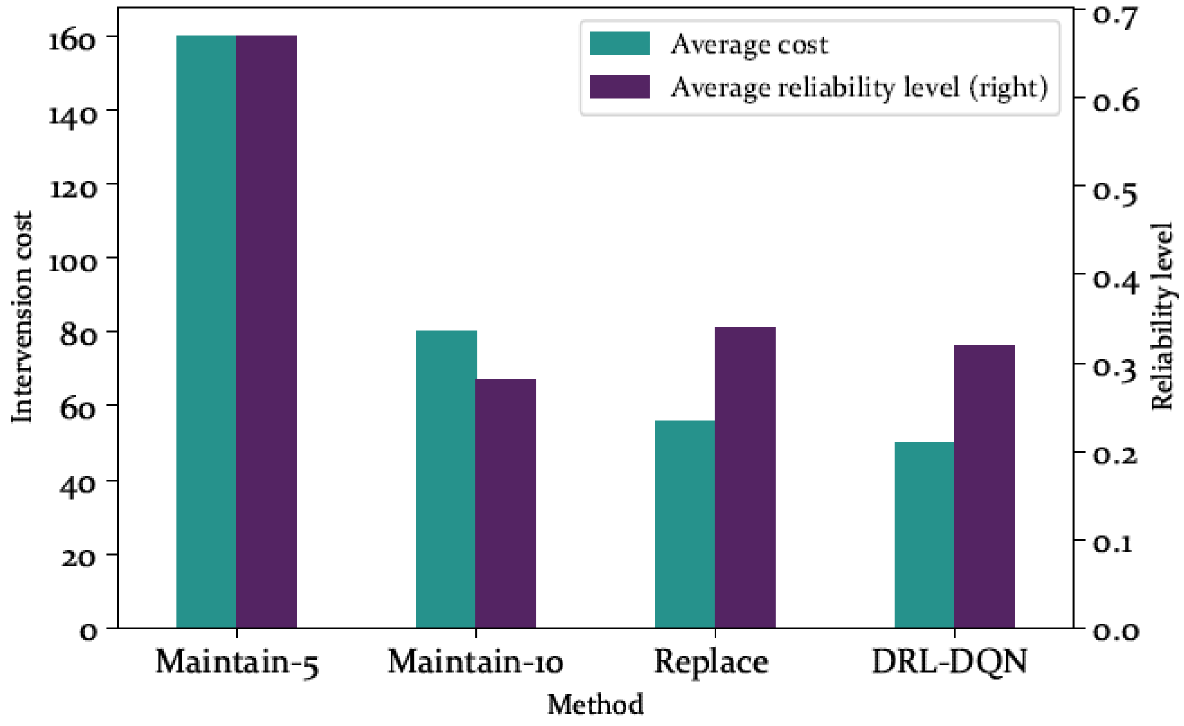


Rehabilitation plan of 16 water pipes for 100 years. Note that the pipes are considered independent from each other in this case study.

Case Study

Comparison with baselines

Comparison of DRL-based rehabilitation method with preventive and corrective approach. The graph shows the average cost (lower values are preferred) and average reliability level (higher values are preferred) for rehabilitation of 16 water pipes for the planning horizon of 100 years.



Method	Cost (Lower is better)	Reliability of network (Higher is better)
Maintain-10	80	0.28
Maintain-5	160	0.67
Replace	56	0.34
DRL	50	0.32

Conclusions and Future steps

- Novel and scalable solution for optimal rehabilitation of water distribution system under economic and reliability requirements

- Adapt a standard DRL framework and DDQN method to learn an optimal policy for the multiple pipes, simultaneously, having distinct physical characteristics.

- Future steps:
  - Comparison with other (policy-based) DRL algorithm
  - Establish baselines for comparison
  - Rehabilitation planning of network segments instead of a single pipe

Thanks!

Got a question?

- ALA Slack channel #paper-45
- Z.bukhsh@tue.nl
- @BukhshZaharah