

강화학습을 이용한 가뭄 상황에서의 댐 방류량 최적화

노하진¹, 임유진²

¹숙명여자대학교 IT공학과 석사과정

²숙명여자대학교 인공지능공학부 교수

hajins@sookmyung.ac.kr, yujin91@sookmyung.ac.kr

Optimization of Dam Discharge in Drought Conditions Using Reinforcement Learning

Hajin Noh¹, Yujin Lim²

¹Dept. of IT Engineering, Sookmyung Women's University

²Div. of Artificial Intelligence Engineering, Sookmyung Women's University

요 약

최근 들어 극심한 가뭄이 지속됨에 따라 댐을 통한 물 수급에 어려움을 겪고 있다. 본 논문에서는 이러한 가뭄 상황에서 댐 자체 방류량 조절을 통해 낭비되고 있는 물을 절약하기 위한 기법을 제안한다. DQN 알고리즘을 사용해 방류량을 최적화하여 목표 저수량 이상의 상태를 60일간 유지하도록 설계하였으며, 해당 알고리즘 내 방류량의 가중치를 변경한 결과를 비교하여 그 성능을 분석하였다.

1. 서론

최근 일부 지역이 심각한 가뭄으로 인해 제한 급수의 위기에 처해있다. 해당 지역의 적극적인 물 절약 홍보에도 불구하고 일정 수준을 회복하지 못하고 있는 실정이다. 향후에는 이러한 재난 상황이 더욱 자주 발생하게 될 것이며, 따라서 피해를 최소화하기 위해서는 다양한 예방 및 대처 방안이 필요하다.

이러한 방안의 하나로 댐 자체 방류량을 절약한다면 제한 급수까지의 기간을 늘릴 수 있을 것으로 기대된다. 물론, 댐에서의 공급량을 줄이는 만큼 관개 및 상수도 시설에서의 수압 조절이나 지자체의 적극적인 안내가 필요할 것이다.

본 논문은 다음과 같이 구성된다. 2장에서는 관련 연구를 소개하고, 3장에서는 목표 및 기준 저수량 설정과 제안하는 알고리즘을 다룬다. 4장은 실험 결과를 분석하며 5장은 결론으로 마무리한다.

2. 관련 연구

[1]에서는 단기 일기 예보와 DQN을 통해 관개 의사 결정을 학습하였으며 일반적인 관개에 비해 뛰어난 물 절약 성능을 보였다. 하지만, 댐의 가뭄은 장기적인 기상 상태의 영향을 받는다는 한계가 있다. [2]는 DQN 기반의 수력 발전 저수지 시스템을 최적화하여 높은 수준의 수력 발전을 제공한다. [3]

에서는 농업 및 상수도 공급, 수력 발전, 홍수 방지 등의 다양한 용도를 고려하였고, 다양한 연속 행동 강화학습 모델을 사용하여 성능을 비교하였다.

본 논문에서는 우리나라 평림댐의 수문 자료를 이용하여 실제 환경과 유사한 기준을 설정하고, 단순화된 댐의 내부 요소만을 환경으로 설정하였다. 의사 결정 과정을 알아보기 쉬운 DQN 학습을 통해 방류량 가중치의 영향을 알아보려고 한다.

3. 시스템 모델

3-1. 목표 및 기준 저수량 설정

학습의 목표는 목표 저수량 이상의 상태를 60일간 유지하는 것이다. 상세한 실험 기준 설정을 위해 물정보포털[4]이 제공하는 2010년 1월 1일부터 2023년 3월 6일까지의 평림댐 수문 자료를 사용하였다.

물정보포털에 따르면, 평림댐은 총 4단계의 대응단계 기준 저수량이 있지만, 보다 섬세한 실험을 위해 '주의'와 '심각'의 기준 저수량 평균값을 '경계' 기준 저수량으로하여 임의 생성하였다. '경계'를 추가한 각 대응단계 기준 저수량은 2023년 2월 6일 기준으로 다음 <표 1>과 같다.

<표 1> 대응단계 기준 저수량(단위: 백만³m)

정상공급환원	관심	주의	경계	심각
4.8	4.3	3.9	2.95	2.0

3-2. 환경 설정

댐의 환경은 0 이상의 값을 가진 저수량, 목표 저수량, 유입량, 방류량으로 이루어진다. 시간 t 에 따라 저수량을 V_t , 유입량을 I , 방류량을 D_t 로 표현한다. 유입량은 변하지 않는 값으로, 시간 t 의 영향을 받지 않는다. 저수량 계산식은 식 (1)과 같다.

$$V_{t+1} = V_t + I - D_t \quad (1)$$

초기 저수량 V_0 는 평립댐 수문 자료 중 저수량이 ‘관심’ 단계(4.3백만 m^3) 이하인 데이터 796개의 저수량 평균인 약 3.36백만 m^3 로 설정하였다. 목표 저수량 V_g 는 현재 가뭄 단계보다 한 단계 위의 기준 저수량으로 설정한다. 따라서, V_0 의 가뭄 단계는 ‘경계’에 해당하며, V_g 는 ‘주의’ 단계의 기준 저수량인 3.9백만 m^3 이 된다. 유입량은 V_0 산정 시 사용한 데이터 중 강우량이 1mm 이하인 데이터 787개의 평균 유입량을 계산한 결과에 24시간을 초로 환산한 값을 곱한 $0.23(m^3/s) \times 86,400(s)$ 로 설정하였다. 방류량 D_t 는 에이전트의 행동에 따라 달라지며 그 값은 <표 2>와 같다. s 는 86,400(s)를 뜻한다.

<표 2> 방류량을 결정하는 행동 집합(단위: m^3)

action 0	action 1	action 2	action 3	action 4	action 5
$0.15 \times s$	$0.175 \times s$	$0.2 \times s$	$0.225 \times s$	$0.3 \times s$	$0.35 \times s$

유입량보다 작은 action 0, 1, 2, 3은 보다 세밀한 절약 행동을 관찰하기 위해 0.25를 기준으로 0.025의 간격으로 설정하였고, action 4, 5는 유입량보다 큰 행동을 선택했음을 확인하기 위한 용도이므로 간격을 0.05로 설정하였다. 종료 조건은 60일 동안 목표 저수량 이상이 유지될 때, 저수율이 약 10%(1백만 m^3) 이하일 때, 또는 365일이 지났을 때로 나뉜다.

3-3. 제안하는 알고리즘

총 보상 r 은 reward와 penalty로 이루어진다. 계산식은 식 (2)와 같다.

$$r = reward - penalty \quad (2)$$

reward의 계산식은 식 (3)과 같다. $0.35 \times 86,400$ 은 행동 집합에서 가장 큰 방류량 값으로, 1.0 이하로 유지하기 위한 값이다. D_t 가 크면 물이 원활하게 공급되므로 V_t 와 D_t 가 클수록 큰 보상을 받는다.

$$reward = \begin{cases} 1.0 + \frac{(V_t - V_g)}{V_g} + \frac{D_t}{0.35 \times 86,400} & (\text{if } V_t \geq V_g) \\ 0 & (\text{else}) \end{cases} \quad (3)$$

penalty의 계산식은 식 (4)와 같다. D_t 가 크면 V_g 에서 멀어지므로 V_t 가 작고 D_t 가 클수록 큰 불

이익을 받는다. w 는 D_t 의 영향력을 조절한다.

$$penalty = \begin{cases} 1.0 + \frac{(V_g - V_t)}{V_g} + w \times \frac{D_t}{0.35 \times 86,400} & (\text{if } V_t < V_g) \\ 0 & (\text{else}) \end{cases} \quad (4)$$

$\frac{V_t - V_g}{V_g}$ (or $\frac{V_g - V_t}{V_g}$) 범위는 reward의 경우 [0, 0.11]이며, penalty의 경우 (0, 0.61)이다. $w \times \frac{D_t}{0.35 \times 86,400}$ 의 범위는 $(0.42 \times w, 1 \times w]$ 이다.

reward의 경우 w 는 1이다. 저수량의 크기보다 방류량을 잘 선택하는 것이 우선이므로 후자에 더 큰 값을 할당하였다.

4. 실험 및 분석

식 (4)에서의 가중치 w 는 0.5, 1, 2, 3의 값을 가질 수 있으며, 최대 에피소드 수는 300이다. <표 3>은 seed별 가중치에 따른 전체 에피소드 중 60일 유지에 성공한 에피소드 개수와 비율을 나타낸다. 비율은 소수점 넷째 자리에서 반올림하였다.

<표 3> seed별 가중치에 따른 목표 달성 에피소드

seed \ w	0.5	1	2	3
12	249(0.83)	285(0.95)	241(0.803)	193(0.643)
34	193(0.643)	151(0.503)	192(0.64)	213(0.71)
56	148(0.493)	186(0.62)	233(0.777)	200(0.667)
78	168(0.56)	218(0.727)	151(0.503)	237(0.79)
90	123(0.41)	153(0.51)	217(0.723)	213(0.71)

w 가 0.5일 때, 성공 비율이 0.65 이상인 경우는 seed가 12일 때 249(0.83)으로 1개이다. w 가 1일 때는 seed 12, 78일 때 285(0.95), 218(0.727)로 2개이며, w 가 2일 때는 seed 12, 56, 90일 때 241(0.803), 233(0.777), 217(0.723) 3개이다. 마지막으로, w 가 3일 때는 seed 12를 제외한 모든 경우이며, 213(0.71), 200(0.667), 237(0.79), 213(0.71)로 4개이다. 또한, 각 seed별 가장 큰 달성률은 w 가 1일 때 285(0.95), w 가 3일 때 213(0.71), w 가 2일 때 233(0.777), w 가 3일 때 237(0.79), w 가 2일 때 217(0.723)으로 w 가 1일 때 1개, w 가 2일 때 2개, w 가 3일 때 2개이다. 이는 penalty에서 방류량의 영향력이 높아질수록 큰 방류량을 피하려는 경향을 보이므로 가중치가 높을수록 seed가 달라져도 높은 달성률을 보일 확률이 높아지며, 달성률도 높은 것을 알 수 있다.

<표 4>는 목표를 달성한 각 에피소드에서 목표 저수량 도달까지 소요 시간을 나타낸다. 목표 달성 비율 차이가 큰 가중치 0.5와 3을 비교하였으며 각 w 와 seed별 w 의 목표 달성 에피소드 개수를 함께

기입하였다. step 1개를 1일로 계산하였다.

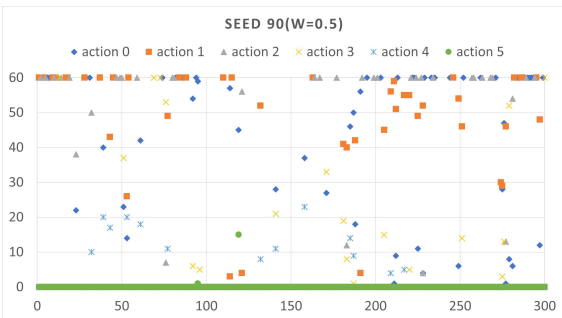
<표 4> 목표 달성 에피소드 중 목표 저수량 도달까지 소요 시간

seed	w	소요 시간		
		78~112	113~207	208~
12	0.5(249)	189(0.759)	22(0.088)	38(0.153)
	3(193)	152(0.787)	20(0.104)	21(0.109)
34	0.5(193)	158(0.819)	16(0.083)	19(0.098)
	3(213)	167(0.784)	21(0.098)	25(0.117)
56	0.5(148)	64(0.432)	38(0.257)	46(0.311)
	3(200)	96(0.48)	64(0.32)	40(0.2)
78	0.5(168)	91(0.542)	52(0.309)	25(0.149)
	3(237)	174(0.734)	49(0.207)	14(0.059)
90	0.5(123)	55(0.447)	40(0.325)	28(0.228)
	3(213)	160(0.751)	30(0.141)	23(0.108)

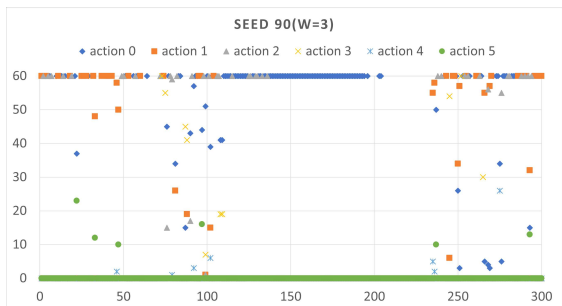
소요 시간에서 78일은 action 0을 78번 선택했다는 의미이고, 113일은 action 1을 113번, 208일은 action 2를 208번 선택한 것이다. 소요 시간이 78일~112일인 비율은 seed 34를 제외한 모든 seed에서 w가 3일 때가 w가 0.5일 때보다 높았다. 이와 비슷하게 소요 시간이 208일 이상인 비율은 seed 34를 제외한 모든 seed에서 w가 0.5일 때가 w가 3일 때보다 높았다. 본 실험 결과는 <표 3>의 결과와 더불어 가중치가 클수록 목표 저수량에 도달하는 속도도 향상된다는 것을 보여준다. 따라서, 더욱 짧은 기간 내로 목표 저수량 달성 및 60일 유지가 가능하다.

하지만, 가중치가 커질수록 단점도 존재한다. (그림 1), (그림 2)는 seed 90에서 에피소드 증가에 따른 V_g 이후 60일간 행동 선택 횟수 분산 정도이다.

(그림 1) seed 90에서 가중치 0.5의 분산 정도



(그림 2) seed 90에서 가중치 3의 분산 정도



(그림 1)은 분산 정도가 큰 것에 비해 (그림 2)는 분산 정도가 비교적 작은 것을 알 수 있다. 이는 가

중치가 클수록 다양한 행동을 선택하지 못한다는 한계점을 나타낸다.

5. 결론

본 연구에서는 가뭄 대처 방안으로 DQN을 이용하여 댐의 방류량을 최적화하는 알고리즘을 제안하였다. 그 중 방류량의 영향을 좌우하는 가중치를 변경해가며 실험을 진행했다. 실험 결과, 가중치가 커질수록 목표 달성 비율이 높아지는 동시에 목표 달성 속도가 빨라지는 경향이 있음을 확인했다. 하지만, 동시에 다양한 선택을 하지 못한다는 문제점이 존재했다. 이러한 결과를 통해 댐 가뭄 해소를 위한 환경 설계에 도움이 될 것으로 기대된다.

추후 연구로는 유입량이 매일 달라지는 실제 환경을 설정하여 그에 따른 방류량을 학습할 것이며, 연속 값을 가지는 강화학습 모델을 사용할 것이다. 또한, 다른 요인들을 추가로 고려하여 실제 상황과 더욱 유사한 환경을 구축하는 연구를 진행할 예정이다.

사사문구

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2021R1F1A1047113).

참고문헌

[1] M. Chen, Y. Cui, X. Wang, H. Xie, F. Liu, T. Luo, S. Zheng and Y. Luo, "A Reinforcement Learning Approach to Irrigation Decision-making for Rice using Weather Forecasts," Agricultural Water Management, vol. 250, Article 106838, May 2021.

[2] X. Wei, M. Fanlin, G. Weisi, L. Xia and F. Guangtao, "Deep Reinforcement Learning for Optimal Hydropower Reservoir Operation," Journal of Water Resources Planning and Management, vol. 147. 04021045. 10.1061/(ASCE)WR.1943-5452.0001409, 2021.

[3] S. T. Sadegh, "Reinforcement Learning Policy Gradient Methods for Reservoir Operation Management and Control," Master's Thesis of Clemson University, 2021.

[4] My Water, https://www.water.or.kr/realtime/sub01/sub01/dam/hydr.do?seq=1408&p_group_seq=1407&menu_mode=2