

제지 공정의 불확실성을 반영하기 위한 Neural SDE 기반 강화학습 연구

유동건, 이상금, 도윤미, 최진영*
국립한밭대학교, *한국전자통신연구원

20211917@edu.hanbat.ac.kr, sangkeum@hanbat.ac.kr, {ydoh, *choij0}@etri.re.kr

Neural SDE-Based Reinforcement Learning for Uncertainty-Aware Papermaking

Donggeon Yu, Sangkeum Lee, Yoonmee Doh, Jinyoung Choi*
Hanbat National Univ., *Electronics and Telecommunications Research Institute

요 약

본 논문은 제지 공정에서 스팀에너지 절감을 위해 스팀 관련 주요 변수를 선별하고, 외부 요인에 의해 발생할 수 있는 이상 패턴과 복합적 이상 징후를 예측하기 위한 Neural SDE 기반 강화학습 모델을 제안한다. 제지 공정은 기상 변화, 지질, 설비 변동 등 생산성과 품질에 직접적인 영향을 미치는 불확실성을 내포한다. 기존 LSTM 기반 모델은 시계열 패턴을 학습하는 데 효과적이거나, 이러한 확률적 변동성을 반영하기 어렵다. 이를 보완하기 위해 본 연구에서는 Neural SDE를 활용하여 공정의 확률적 특성을 반영하였으며, 다중 보상 함수를 통해 스팀에너지 효율을 포함한 다양한 목표를 동시에 고려하였다. 실험 결과, SDE와 LSTM을 결합한 모델의 예측률이 약 5% 증가하였다. 제안한 모델을 통해 제지 공정의 불확실성을 반영하면서 공정 최적화에 기여할 수 있는 가능성을 확인하였다.

I. 서 론

통계에 따르면 2020년 국내 제지 산업 열에너지 소비의 약 11.9%를 차지한다[1]. 국가 탄소중립 목표에 부합하려면, 공정 단계에서의 효율 개선이 가장 직접적인 감축 수단이다[2]. 제지공정의 경우, 초지공정에서 60% 이상의 스팀이 사용된다. 또한, 지질로 인한 생산 지중 변경, 계절 등 외부 요인으로 인해 운전 조건이 변경된다. 따라서 본 연구에서는 스팀사용량을 최적화하기 위해 스팀과 관련된 변수를 선별하고, 운전에 따른 편차를 최소화하기 위한 강화학습 모델을 개발하였다. 강화학습에서 Agent는 Environment와 상호작용하며 정책을 학습한다. Environment는 물리 방정식 또는 인공지능으로 구현될 수 있다. 그러나 단순 물리 방정식이나 단일 인공지능은 제지공정에서 발생하는 급격한 변동성, 제품 생산 주기, 노이즈 등 실제 공정 환경을 반영하기 어렵다. 단순 모델의 일반화를 배제하면서, 품질, 에너지 효율과 같이 다중 목표를 동시에 최적화하는 공정 맞춤형 제어 방법이 필요하다. 본 연구는 LSTM 기반 Environment에 Neural Stochastic Differential Equation (SDE)을 추가하여 실측값에 근접한 값을 예측하는 모델을 개발하였다. 또한, 기존 작업자 운전 조건 대비 제안한 방법 운전 조건의 성능을 점수화하여 비교하였다.

II. 본론

2.1 Neural SDE+ LSTM 강화학습 개념도

그림 1은 Neural SDE이 추가된 LSTM 기반(이하 SDE+ LSTM) 강화학습 모델의 개념도이다. LSTM이 1차로 예측값을 출력하고, Neural SDE을 거쳐 최종 예측값을 출력한다. 공정 환경에서는 외부 요인들이 변함에 따라 실제 운영의 데이터가 달라지는 데이터 드리프트 현상이 존재한다. 이를 보완하기 위해 실측값과 LSTM의 예측값 차이를 학습하여 오차를 보정하고, 노이즈를 추가하여 강건

함을 유지하면서 실측값에 더 근접한 학습이 가능한 강화학습 모델을 개발하였다.

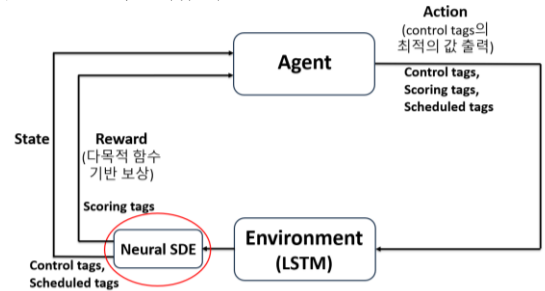


그림 1. Neural SDE+ LSTM 강화학습 개념도

$$dX_t = \underbrace{\mu(t, X_t)dt}_{\text{drift}} + \underbrace{\sigma_{\theta}(t, X_t)dW_t}_{\text{diffusion}} \quad (1)$$

$$MOF = \sum_{j=1}^3 a_j \cdot \frac{1}{|T_j|} \sum_{t \in T_j} \left(\sum_{i=1}^n W_i \cdot e^{-\left(\frac{x_i(t) - x_i^{ref}}{x_i^{tol}} \right)^2} \right) \quad (2)$$

표 1. 수식 기호별 의미

수식 기호	의미
$x_i(t)$	i 번째 x 의 시간 구간
x_i^{ref}	i 번째 x 의 기준값
x_i^{tol}	i 번째 x 의 허용 편차
W_i	i 번째 중요도 가중치
a_j	j 번째 시간 구간의 가중치
T_j	j 번째 시간 구간

Neural SDE는 확률미분방정식을 기반으로 한 데이터 생성 모델이다[3]. 수식 (1)은 예측값을 계산하기 위한 공식이다. *drift*는 다음 시점 예측값의 변화 방향을 나타내며, *diffusion*은 현재 시점과 값들을 기반으로 노이즈의 강도와 방향을 결정하여 예측값으로 입력된다. Reward는 Agent가 최대화하도록 학습되는 대상이며, 수식 (2)의 다

목적 함수(MOF: Multi objective Function)가 Reward로 사용된다. MOF의 점수는 0과 1 사이이며 기준값에 가까울수록 1에 가까워지는 함수이다. 기준값은 생산 지중의 작업량의 평균으로 수식의 지수항은 측정값이 기준값에서 벗어날수록 가우시안 형태의 페널티를 적용하여, 안정적인 범위 내에서 설비 조건이 변경 가능하도록 유도한다.

2.2 데이터 전처리 및 특징

본 연구에서는 국내 제지 공정에서 2022년 5개월간 1분 간격으로 수집된 시계열 데이터를 사용하였다. 현장 작업자 경험과 상관관계 분석을 바탕으로 스팀사용량에 영향을 미치고 결측치 비율이 낮은 변수를 선별하여 사용하였다. 선정된 변수의 이상치는 Hampel 필터로 전처리하였고, 결측치는 평균 보간법을 적용하였다[4]. 선정된 변수는 3가지의 카테고리로 분류하였으며, 세부 내용은 표 2에 나타내었다.

표 2. 제지공정의 주요 변수와 카테고리

카테고리	변수	설명
Control tags	스팀	스팀사용량
Scoring tags	평량1	전건조부 평량 측정값(g/m^2)
	평량2	후건조부 평량 측정값(g/m^2)
	수분량	전건조부 수분량 측정값(%)
	전건조기	전건조부 스팀 압력 측정값(kg/cm^2)
	후건조기	후건조부 스팀 압력 측정값(kg/cm^2)
	릴_속도	릴 속도 측정값(m/min)
	유량	유량 측정값(m^3/h)
Scheduled tags	평량	목표 종이 평량(g/m^2)
	LOT_진행	작업 진행률(%)
	제조_shift	제조번호 변화 이벤트
	속도_계산	계산된 리 속도(m/min)
	속도_표준	릴 목표 속도(m/min)
	지중_그룹	유사한 지중끼리 클러스터링(5개)

- Control tags: Agent는 다중 변수값을 고려해 Reward를 최대화하는 스팀사용량을 출력한다.
- Scoring tags: Reward 계산에 사용되는 다중 변수들로, 종이 품질과 에너지 사용에 직접적인 영향을 미치는 핵심 변수들로 구성된다. 본 연구에서는 상관관계가 높은 7개 변수가 선정되었다.
- Scheduled tags: Scoring tags와 관련된 사전정보 다중 변수들로, 미래값이 LSTM의 입력으로 사용된다. '지중_그룹'은 유사한 지중끼리 클러스터링하여 5개의 그룹으로 분류된 변수를 의미한다.

2.3 예측 결과 및 성능 비교

표 3. LSTM과 SDE+ LSTM의 예측 결과

지표	LSTM	SDE+ LSTM	차이
R^2	0.8968	0.9495	+0.0527
RMSE	0.3166	0.2215	-0.0951
EMD	0.8900	0.9062	+0.0162

표 3의 지표를 살펴보면, SDE+ LSTM의 R^2 값이 1의 근사치인 것을 통해 해당 모델에서 예측한 값의 설명력이 더욱 높으며, RMSE 값이 감소한 것을 통해 기존 LSTM보다 에러를 감소시켰다. 모델의 개별 예측값 정확도를 확인하기 위한 두 지표 외 예측 분포가 실제 데이터 분포와 얼마나 유사한지 확인하기 위해 EMD(Earth-Mover Distance)를 수행하였고, 값이 상승한 것을 통해 불확실성

이 반영되었음을 확인하였다. 이는 논문에서 제안한 방법이 기존 모델보다 예측력이 높고, 실측 데이터의 변동성과 노이즈 패턴의 재현성이 높은 것을 의미한다. 그림 2는 실측값, LSTM의 예측값, SDE+ LSTM의 예측값을 출력한 예시이다. LSTM이 학습한 것보다 SDE+ LSTM이 실측 데이터에 더 가깝게 예측한 것을 확인하였다. 표 4는 기존 운전 조건일 때와 강화학습 적용 운전조건의 다목적 함수를 비교한 표이다. 이는 강화학습 적용 시, 스팀 관련 주요 변수들의 편차가 감소했음을 의미한다.

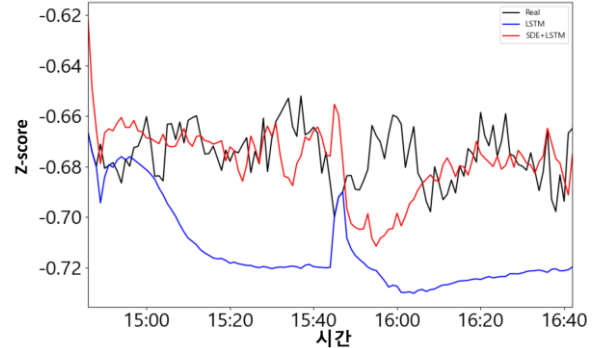


그림 2. 모델별 예측값과 실측값 예시

표 4. 기존 운전조건과 Neural SDE+ LSTM 강화학습의 MOF 점수 결과

성능지표	기존 운전조건	RL 운전조건	차이
MOF	0.5355	0.5913	+ 0.0452

III. 결론

본 연구에서는 제지공정의 불확실성을 줄이기 위한 강화학습 연구를 수행하였다. 공정 데이터의 변동성과 노이즈를 반영하지 못하는 단일 LSTM을 보완하기 위해 Neural SDE를 활용한 강화학습 모델을 개발하였다. SDE+ LSTM의 예측력과 재현성이 기존 단일 모델보다 약 5% 증가한 것을 확인하였다. 향후 에너지데이터를 추가 분석하여 에너지 효율을 분석하고, 다양한 강화학습 모델과의 성능 비교를 통해 모델 검증 및 고도화 연구를 진행할 예정이다.

ACKNOWLEDGMENT

본 연구는 산업통상자원부(MOTIE)와 한국에너지기술연구원(KETEP)의 지원을 받아 수행한 연구 과제입니다. (No. RS-2020-KP002211)

참 고 문 헌

- [1] 한국에너지공단, “EG-TIPS 에너지온실가스종합정보 플랫폼: 국가 인벤토리 기준 산업부문 세부업종별 배출량(2020년),” EG-TIPS. [Online]. Available: https://tips.energy.or.kr/carbon/Ggas_industry.do. Accessed: 2025-08-19.
- [2] Republic of Korea, “Framework Act on Carbon Neutrality and Green Growth for Coping with Climate Crisis, Art. 7 (National Vision and Strategy),” English translation, Korea Legislation Research Institute (KLRI). Accessed: Aug. 19, 2025.
- [3] B. Tzen and M. Raginsky, “Neural stochastic differential equations: Deep latent Gaussian models in the diffusion limit,” arXiv preprint arXiv:1905.09883, 2019.
- [4] R. K. Pearson et al., “Generalized Hampel filters,” EURASIP Journal on Advances in Signal Processing, 2016:87, 2016.