제 28 권 제 1 호
Vol. 28 No. 1

2026

제 28 회

한국 소프트웨어공학 학술대회 논문집

Proceedings of the 28th Korea Conference on Software Engineering (KCSE 2026)

- 일시: 2026년 2월 4일(수) ~ 2월 6일(금)
- 장소: UNIST(울산과학기술원) 114동 경영관

주최: 한국정보과학회, 한국정보처리학회
주관: 한국정보과학회 소프트웨어공학 소사이어티
한국정보처리학회 소프트웨어공학연구회

후원:

세션 D1. SW 테스팅 II

- 2026년 02월 06일 (금) 오전 10:15-11:45 / 102호
- 좌장: 손정주 교수(경북대)

[초청발표] How Effective are Large Language Models in Generating Software Specifications?,
류병우, 김미정(UNIST), Danning Xie, Nan Jiang, Lin Tan, Xiangyu Zhang(Purdue University), The 32nd IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER 2025) 국제학술대회 발표 논문
발표자: 류병우(UNIST)

질의-응답 프롬프트 기반 실용적인 테스트 오라클 생성 (일반논문)
정지나, 김윤호(한양대)

상태 공간 모델 기반 오프라인 강화학습의 강건성 테스팅 (단편논문)
한태현, 김장환, 김영철(홍익대)

모델체킹을 위한 강화학습 기반 휴리스틱 학습 (단편논문)
강혜윤, 손병호, 배경민(POSTECH)

[초청발표] TopSeed: Learning Seed Selection Strategies for Symbolic Execution from Scratch,
이재혁, 차수영(성균관대), IEEE/ACM 47th International Conference on Software Engineering (ICSE 2025) 국제학술대회 발표 논문
발표자: 이재혁(성균관대)

세션 D2. SW 디버깅

- 2026년 02월 06일 (금) 오전 10:15-11:45 / 111호
- 좌장: 지은경 교수(KAIST)

[초청발표] Collaboration failure analysis in cyber-physical system-of-systems using context fuzzy clustering, 현상원(University of Adelaide), 지은경, 배두환(KAIST), Empirical Software Engineering(EMSE 2025) 국제저널 발표 논문
발표자: 현상원 (University of Adelaide)

딥러닝 기반 국방 SW 결함위치추정을 위한 변이 기반 데이터셋 구축의 체계적 연구 (단편논문)
양희찬, 이아청(KAIST), 조규태(LIGNex1), 김문주(KAIST/브이플러스랩)

이슈 설명과 심볼 수준의 목표를 활용한 분류 기반 결함 위치 식별 (단편논문)
Aslan Safarovich Abdinabiev, 홍수지, 이병정(서울시립대)

실행 컨텍스트 정합성에 기반한 LLM 결함 위치 추정 결과의 안정화 기법 (일반논문)
남규민, 최서진, 양근석(한경국립대)

Diff-only 환경에서 단일 에이전트와 다중 에이전트 기반 커밋 메시지 생성의 비교 분석 (일반논문)
안도경, 양근석(한경국립대)

단편 논문

상태 공간 모델 기반 오프라인 강화학습의 강건성 테스팅

한태현, 김장환, 김영철
홍익대학교 소프트웨어공학연구실
taehyun3172@g.hongik.ac.kr, lentoconstante@hongik.ac.kr, bob@hongik.ac.kr

Robustness Testing of Offline Reinforcement Learning based on State Space Models

Taehyun Han, Janghwan Kim, R. Young Chul Kim
Software Engineering Laboratory, Hongik University

요 약

최근 오프라인 강화학습이 실제 환경에 도입됨에 따라, 예측 불가능한 외부 노이즈에 대한 안전성 및 강건성 검증이 필수적이다. 그러나 기존 무작위 노이즈 방식은 취약 시점을 식별하지 못해 비효율적이며, 적대적 공격 기법은 높은 연산 비용으로 인해 실시간 검증에 적용하기 어렵다. 이러한 문제를 해결하기 위해, 상태 공간 모델 기반 Mamba 아키텍처를 활용한 효율적인 블랙박스 테스팅 기법을 제안한다. 본 연구는 노이즈의 주입 시점이 시스템의 생존 및 강건성에 결정적임을 실험적으로 입증하였다. 이를 통해 제안하는 Mamba 기반 노이즈 주입 기법이 효율적인 블랙박스 테스팅 기법으로서 활용될 것을 기대한다.

1. 서론

최근 로보틱스와 같은 복잡한 제어 문제에서 사전에 수집된 데이터를 활용하는 오프라인 강화학습이 핵심 기술로 부상하고 있다. 특히 Decision Transformer (DT)는 강화학습을 시퀀스 모델링으로 재해석하여 탁월한 성과를 거두었다 [1]. 그러나 이러한 모델이 안전 필수 환경에 배포되기 위해서는 예측 불가능한 외부 노이즈에 대한 강건성 검증이 선행되어야 한다 [2].

현재 강건성 평가 방법론은 크게 무작위 노이즈 주입과 적대적 공격 기법으로 나뉜다. 무작위 방식은 구현이 간단하나, 에이전트의 실패를 유발하는 결정적 원인을 파악하기 어렵고 효율이 낮다 [3]. 반면, 적대적 공격 기법들은 탐지 성능은 우수하나, 높은 연산 비용과 학습 복잡도로 인해 실시간 진단 도구로 활용하기에는 제약이 존재한다 [4].

본 논문에서는 이러한 한계를 극복하기 위해 상태 공간 모델(SSM)인 Mamba 아키텍처를 활용한 효율적인 강건성 테스팅 프레임워크를 제안한다 [5,6]. 우리는 Mamba 가 환경의 동역학을 학습하는 과정에서 내부 파라미터인 델타 (Δ) 값이 급증하는 구간이 곧 정보량이 높고 노이즈에 취약한 불안정한 상태라는 가설을 제시한다. 제안 기법은 Mamba 의 델타 파라미터 지표를 모니터링하여 불안정한 시점의 상태를 식별하고, 선택적으로 노이즈를 주입하여 모델의 취약점을 효율적으로 진단한다.

2. 상태 공간 모델 기반 강건성 테스팅 메커니즘

SSM 은 연속적 시간의 시스템을 이산화하며, 이때 이산화 파라미터 델타 (Δ) 는 데이터를 얼마나 반영할지 결정하는 역할을 한다 [5]. 기존 SSM 과 달리 Mamba 는 입력 x에 따라 Δ가 동적으로 변하는 선택적 메커니즘을 도입하였다 [6].

$$\Delta_t = Softplus(W_\Delta x_t + b_\Delta) \quad (1)$$

수식 1 과 같이 Δ_t 는 고정된 값이 아닌 매 시점 입력 데이터와 학습된 파라미터의 조합에 의해 결정되는 동적 행렬이다. 본 연구는 이러한 Δ가 급격히 변하는 시점이 곧 모델이 예측하기 어려운 불안정 상태라는 가설에 기반한다.

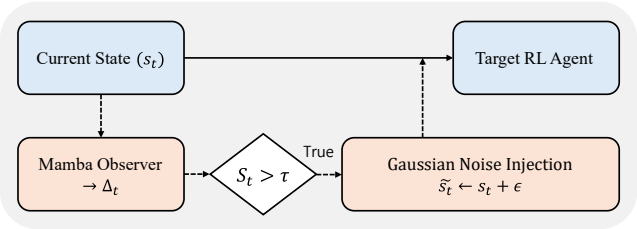

그림 1. 제안하는 Mamba 기반 강건성 테스트 프레임워크

그림 1 은 제안하는 전체 프레임워크의 구조를 보여준다. 제안 기법은 에이전트의 정책과 무관하게 환경의 특성을 분석하기 위해 Mamba 아키텍처를 관측 모델로 활용한다. 오프라인 데이터셋을 사용하여 관측 모델은 현재 시점까지의 상태 $s_{1:t}$를 입력 받아 다음 상태 s_{t+1}을 예측하도록 학습한다 [8,9]. 학습된 관측 모델은 테스팅 단계에서 실시간으로 입력되는 상태 s_t에 대해 내부 파라미터 Δ_t를 산출한다. 이를 지표로서 활용하기 위하여 다음 수식 2 와 같이 L2-Norm 을 적용하여 단일 스칼라 값인 중요도 점수 (S_t)로 변환한다.

$$S_t = ||\Delta_t||_2 = \sqrt{\sum_{i=1}^{D} (\Delta_{t,i})^2} \quad (2)$$

여기서 S_t 값이 클수록 현재 상태가 미래 예측에 중요한 정보를 담고 있음을 의미한다. 제안 기법은 S_t 가 사전에 설정된 임계치 (τ) 를 초과하는 경우 시스템의 취약 시점으로 판단하여 가우시안 노이즈를 주입하고 강건성을 검증한다.

3. 상태 공간 모델 기반 강건성 테스팅 메커니즘 적용 사례

제안 기법의 유효성을 검증하기 위해 D4RL 벤치마크의 MuJoCo 환경(Hopper, Walker2d, HalfCheetah)을 사용하였으며 검증 대상 모델로는 Medium 데이터셋으로 학습된 DT 모델을 사용하였다 [1,7]. 이때 Mamba 관측 모델의 하이퍼 파라미터는 임베딩 차원 $D = 64$, 상태 차원 $N = 16$ 으로 설정하여 대상 모델과 동일한 데이터셋으로 학습시켰다. 노이즈 주입 임계치 (τ)는 사전 에피소드 수행을 통해 수집된 Δ값의 분포의 상위 10%로 결정하였다.

3.1 실험 결과 및 분석

표 1. 환경 별 성능 하락률 결과

Noise		Performance Drop		
Scale	Method	HalfCheetah	Hopper	Walker2d
0.1	Random	6.0%	6.5%	8.3%
	Mamba	9.5%	24.3%	8.3%
0.2	Random	8.1%	14.3%	11.9%
	Mamba	17.8%	39.8%	5.5%
0.3	Random	10.8%	16.6%	14.0%
	Mamba	34.6%	53.5%	19.9%
0.4	Random	10.0%	23.6%	29.9%
	Mamba	26.3%	60.6%	40.2%
0.5	Random	10.0%	22.8%	22.9%
	Mamba	37.8%	61.7%	46.8%

표 1 은 각 환경에서의 제안 기법 적용 결과이다. 실험 결과, 로봇의 물리적 구조에 따라 제안 기법의 효과가 뚜렷한 차이를 보였다. 구조적으로 가장 불안정한 Hopper 에서는 제안 기법의 효용성이 극대화되었다. 최소 노이즈 스케일만으로도 무작위 방식이 최대 스케일을 가했을 때보다 더 큰 성능 하락을 유발하며 취약 시점의 중요성을 입증하였다. 이족 보행 로봇인 Walker2d 에서는 노이즈 강도가 높아질수록 두 방식 간의 격차가 벌어졌다. 최대 스케일에서 제안 기법은 무작위 방식 대비 2 배 이상의 성능 저하를 기록하였다. 마지막으로 가장 안정적인 HalfCheetah 에서도 무작위 방식은 성능 저하가 미미했던 반면, 제안 기법은 약 4 배에 달하는 효과를 보였다. 종합적으로 제안 기법은 모든 환경에서 무작위 방식 대비 월등한 성능 하락을 유도하여, 효율적인 강건성 테스팅 도구임을 증명하였다.

3.2 선택적 메커니즘과 불안정 상태의 상관관계

본 연구는 Mamba 의 내부 파라미터인 델타(Δ)가 급증하는 구간이 곧 에이전트의 불안정 상태라는 가설을 검증하였다. 선택적 메커니즘에서 Δ 는 정보 반영률을 결정하므로, 높은 중요도 점수(S_t)는 해당 상태가 미래 예측에 핵심적인 정보를 담고 있음을 의미한다. 실험 결과, Δ 값이 높은 시점에 노이즈를 주입했을 때 무작위 방식 대비 치명적인 성능 하락이 발생하여 이 가설의 유효성이 입증되었다. 이는 복잡한 연산이나 모델 내부의 기울기 정보 없이도, 관측 모델의 내부 지표만으로 시스템의 취약 시점을 효율적으로 식별할 수 있음을 시사한다.

4. 결론

본 논문에서는 오프라인 강화학습 모델의 강건성을 효율적으로 평가하기 위해, SSM 기반 Mamba 아키텍처를 활용한 테스팅 프레임워크를 제안하였다. 우리는 Mamba 의 내부 지표인 델타(Δ)를 활용하여 에이전트가 외부 노이즈에 민감하게 반응하는 불안정한 상태를 식별하고, 해당 시점에 선별적으로 노이즈를 주입하였다. 실험 결과, 제안 기법은 무작위 방식과 동일한 빈도의 노이즈 주입만으로도 최대 4 배 이상의 보상 하락을 유도하였다. 이는 제안 기법이 에이전트의 제어 실패를 유발하는 결정적 순간을 효과적으로 식별했음을 시사하며, 효율적인 블랙박스 강건성 테스팅 도구로서의 활용 가능성을 보여준다. 본 연구는 무작위 방식과의 비교를 통해 제안 기법의 기초적인 타당성을 확인하였으며, 향후 연구에서는 단순 휴리스틱 및 기존 적대적 공격 기법과의 연산 비용 및 탐지 성능에 대한 정량적 비교를 수행하여 Mamba 내부 지표의 고유한 효용성을 더욱 정밀하게 검증하고자 한다.

감사의 글

본 연구는 한국연구재단의 4 단계 두뇌한국 21 사업(과제명: 초분산 자율 컴퓨팅 서비스 기술 연구팀, 과제번호: 202003520005)의 지원을 받아 수행된 연구임.

5. 참고문헌

[1] L. Chen et al., Decision transformer: Reinforcement learning via sequence modeling, *Advances in Neural Information Processing Systems*, Vol. 34, pp. 15084-15097, 2021.

[2] S. Huang et al., Adversarial attacks on neural network policies, *arXiv preprint arXiv:1702.02284*, 2017.

[3] Y. Wang et al., A systematic review of fuzzing based on machine learning techniques, *PLoS One*, Vol. 15, No. 8, e0237749, 2020.

[4] J. Sun et al., Stealthy and efficient adversarial attacks against deep reinforcement learning, *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34, No. 04, 2020.

[5] A. Gu et al., Efficiently modeling long sequences with structured state spaces, International Conference on Learning Representations (ICLR), 2022.

[6] A. Gu and T. Dao, Mamba: Linear-time sequence modeling with selective state spaces, *First Conference on Language Modeling*, 2024.

[7] J. Fu et al., D4rl: Datasets for deep data-driven reinforcement learning, *arXiv preprint arXiv:2004.07219*, 2020.