

소프트웨어 중심 선박 환경에서의 희귀 데이터 기반 자율 진화 학습을 위한 선박 간 데이터 교환 메커니즘

임민중, 김기현, 나영진
동국대학교

minjoong@dongguk.edu, rlsrlgus123@naver.com, dudwls976@naver.com

Ship-to-Ship Data Exchange Mechanisms for Rare Event-Driven Learning in Software-Defined Ship Environments

Minjoong Rim, Ki-Hyeon Kim, Young-Jin Na
Dongguk University

요 약

본 연구는 소프트웨어 중심 선박 환경에서 희귀 데이터 기반 자율 진화형 학습을 지원하기 위한 선박 간 데이터 교환 알고리즘을 제안한다. 소프트웨어 중심 선박에서는 지속적 학습을 통해 선박의 기능과 성능을 자율적으로 고도화할 수 있으나, 악천후, 이상 징후, 고장 등 드문 사건에 대한 데이터 수집 및 학습은 여전히 큰 도전 과제이다. 이를 해결하기 위해 본 연구는 선박이 항해 중 수집하는 데이터를 희귀성, 예측 오차, 위험성 기준에 따라 점수화하고, 중요도에 따라 클래스로 구분하여 저장하는 전략을 설계하였다. 저장 용량이 제한된 상황에서 상위 클래스 데이터를 우선 저장하며, 선박 간 통신이 가능할 때, 저장 여유 공간, 남은 항해 시간, 데이터 중요도, 데이터 발생 분포 등을 고려해 데이터를 교환하는 알고리즘을 제안하였다. 시뮬레이션 결과, 제안하는 교환 전략은 선박 간 통신 확률이 증가할수록 상위 클래스 데이터의 중앙 서버 도달 비율을 향상시킨다.

I. 서론

소프트웨어 중심 선박은 기존 하드웨어 기반의 선박 설계 및 운용 방식에서 벗어나, 소프트웨어를 중심으로 선박의 제어, 운용, 유지보수 및 업그레이드를 수행할 수 있도록 설계된 차세대 선박 기술이다[1,2]. 소프트웨어 중심 선박은 다양한 해양 환경에서 발생하는 데이터를 지속적으로 수집하고, 수집된 데이터를 기반으로 인공지능(AI) 모델을 지속적으로 학습시킴으로써 선박의 자율적 기능 진화와 운영 최적화를 가능하게 한다.



그림 1. 소프트웨어 중심 선박의 선순환 학습 구조

기존의 스마트 선박 시스템은 사전에 수집된 데이터로 학습한 알고리즘을 선박에 탑재하여 고정적으로 운용한다. 그러나 이러한 방식은 실제 운항 중 발생하는 환경 변화, 기상 이변, 고장 징후와 같은 드문 사건들에 대해 적응 능력이 제한된다. 특히, 악천후, 센서 오작동, 긴급 회피 기동, 예기치 못한 항로 변경 등의 상황은 빈도는 낮으나 위험도가 높기 때문에, 이를 효과적으로

학습하고 관리하는 것이 선박의 자율성과 안전성을 향상시키는 데 필수적이다.

[3]에서는 항해 중 선박 간 사이드링크(sidelink)를 활용하여 데이터를 공유하고, 통신 환경이 양호한 해역에 진입했을 때 수집된 데이터를 중앙 서버로 전송하는 방식을 제안하였다. 본 논문은 이를 기반으로 더욱 발전된 모델을 제안하고, 선박 간 데이터 교환이 가능한 환경에서 데이터의 우선순위 평가와 교환 알고리즘을 설계한다. 제안하는 알고리즘은 높은 가치의 데이터를 효율적으로 관리하고 공유하여 희귀 데이터의 손실을 최소화하며, AI 모델의 학습 효율을 높이는 것을 목표로 한다.

II. 클래스 기반 희귀 데이터 평가 및 저장 전략

각 선박은 항해 중 다양한 센서(카메라, 라이다, 레이더, GPS, IMU, 환경 센서 등)를 통해 대량의 데이터를 지속적으로 수집하며, AI 학습 관점에서 중요한 데이터를 선별한다. 본 연구에서는 다음과 같은 기준을 종합적으로 고려하여 데이터의 상대적 가치를 평가하고, 이 가치를 기반으로 데이터를 클래스로 구분한다:

- 희귀성(Rarity): 폭풍우, 짙은 안개, 높은 파도, 해양 구조물 근접, 항구 진입, 이상 징후, 고장 징후 등 비정상적이고 드문 상황
- 예측오차(Prediction Error): AI 예측과 실제 값 차이, 센서 간 판단 불일치, 모델의 불확실성
- 위험성(Risk Level): 충돌 위험, 장애물과 근접, 불법/위험 해역 진입, 복잡 경로, 급가속/급선회 등의 위험 상황

수집된 데이터는 위 요소를 기반으로 수치화되며, 데이터의 상대적 중요도에 따라 N_{class} 개의 클래스로 나눈다 (낮은 번호가 높은 중요도). 본 연구는 상위 클래스 데이터를 가급적 손실 없이 중앙 서버로 수집하는 것을 목표로 하며 다음과 같이 동작한다.

for k 는 1 부터 N_{class} 까지
 클래스 k 미만의 데이터의 손실을 희생하지 않으면서
 클래스 k 데이터 손실 최소화 수행

각 선박은 저장장치 용량이 제한되어 있으며, 전체 수집 데이터 중 일부만 저장 가능하다. 저장된 데이터는 클래스 기반 우선순위 정책에 따라 운영되며, 새로운 상위 클래스의 데이터가 발생하면 하위 클래스의 데이터를 제거하여 교체한다. T_{unit} 의 시간 단위마다 전체 선박 수 N_{ship} 중 $P_{comm}N_{ship}/2$ 쌍의 선박이 서로 통신할 수 있는 범위에 진입한다고 가정한다. 이 때 선박 n 에 대해 저장된 데이터 양은 다음과 같이 표시한다.

- 시간 단위: T_{unit}
- 항해 시간: $T_{trip}(n)$
- 현재 경과 시간: $T_{current}(n)$
- 클래스의 수: N_{class}
- 클래스 k 의 발생 확률: $P_{class}(n, k)$
- 데이터 블록 단위: D_{block}
- 저장 용량: $D_{size}(n)$
- 현재 저장된 클래스 k 의 데이터 양: $D_{class}(n, k)$

어떤 k_0 값에 대해서 상대적으로 중요성이 떨어지는 클래스 k_0 초과 데이터를 무시한다고 할 때, 항해 종료 시의 예상 저장 여유 공간 $D_{margin}(n, k_0)$ 은 다음과 같이 계산한다.

$$D_{margin}(n, k_0) = D_{size}(n) - \sum_{j=1}^{k_0} D_{class}(n, j) - (T_{trip}(n) - T_{current}(n)) \sum_{j=1}^{k_0} P_{class}(n, j) \quad (1)$$

선박 간 통신이 가능할 때, 통신범위 내 두 선박은 서로의 여유 공간을 비교하고 더 많은 여유 공간을 가진 선박이 상위 클래스 데이터를 수용하도록 하여, 회귀 데이터의 중앙 서버 도달 확률을 높인다. 구체적인 알고리즘은 다음과 같다.

for k 는 1 부터 N_{class} 까지
 while 더 이상 교환이 이루어지지 못할 때까지
 다음 조건 확인
 한 선박 $n1$ 이 다른 선박 $n2$ 보다 여유 공간 보유

$$D_{margin}(n1, k) > D_{margin}(n2, k) + 1 \quad (2)$$

선박 $n1$ 이 추가 저장 가능

$$\sum_{j=1}^k D_{class}(n1, j) < D_{size}(n1) \quad (3)$$

선박 $n2$ 에 클래스 k 데이터가 존재

$$D_{class}(n2, k) > 0 \quad (4)$$

모두 만족 시 선박 $n2$ 의 클래스 k 데이터와 선박 $n1$ 에 저장된 최하 클래스 데이터 교환

III. 시뮬레이션 및 성능 분석

실험에서는 선박 간 통신 확률 P_{comm} 에 따라 클래스 1~ k 데이터가 중앙 서버에 도달하지 못하고 손실될 확률을 측정한다. P_{comm} 이 0 인 경우는 선박 간 통신 가능 여부와 상관 없이 데이터 교환을 하지 않는 경우라고 볼 수 있다. 실험 변수는 다음과 같다.

- 선박 수 N_{ship} : 100 (한 선박의 항해가 끝나면 바로 다른 선박 항해 시작)

- 시뮬레이션 단위: T_{unit}
- 항해 시간 분포: 1/2 확률로 50 또는 100 T_{unit} 시간
- 항해 시작 시간: 랜덤 시작
- 저장 용량 $D_{storage}(n)$: 1/2 확률로 5 또는 20 블록
- 데이터 클래스 수 N_{class} : 10
- 클래스 발생 확률 $P_{class}(n, k)$: 1, 2, 3, 4, 15, 15, 15, 15, 15, 15%

그림 2 에서 선박 간 통신 확률이 증가함에 따라 클래스 1~ k 데이터의 손실률이 감소함을 볼 수 있다.

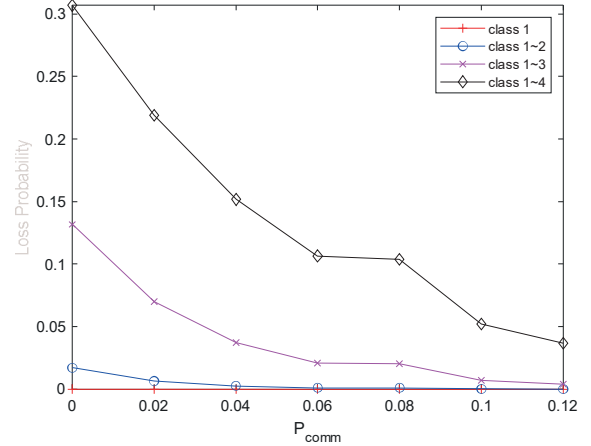


그림 2. P_{comm} 에 따른 중요 클래스 데이터 손실률

VI. 결론

본 연구는 소프트웨어 중심 선박 환경에서, 제한된 저장공간 및 통신환경 하에서도 회귀 데이터를 효율적으로 수집하고 중앙 서버로 전달하는 알고리즘을 제안하였다. 클래스 기반 저장 전략과 여유 공간 예측을 이용한 선박 간 데이터 교환 메커니즘을 통해 AI 학습에 유의미한 데이터를 확보할 수 있으며, 이는 소프트웨어 중심 선박의 자율 진화형 AI 아키텍처 실현에 중요한 기반이 될 것이다. 향후 연구 방향은 다음과 같다.

- 연합학습 연계: 선박 간 모델 파라미터 공유 및 병렬 학습 방안 연계
- 위성 연계: LEO 기반 통신 시스템 통합 효과 분석
- 다대다 통신 기반 최적 매칭 알고리즘: 선박 간 네트워크 모델을 기반으로 한 동적 교환 전략
- 보안과 프라이버시: 선박 간 데이터 공유 시 보안 및 개인정보 보호 방안 고려
- 강화학습 기반: 실시간으로 데이터의 중요도 및 알고리즘을 조정하는 적응형 클래스 정책 도입

ACKNOWLEDGMENT

본 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-학석사연계 ICT 핵심인재양성 지원을 받아 수행된 연구임 (IITP-2024-00436744)

참 고 문 헌

- [1] C. Zhu, W. Zhang, Y. Chiang, N. Ye, L. Du, and J. An, "Software-Defined Maritime Fog Computing: Architecture, Advantages, and Feasibility," IEEE Network, vol. 36, no. 2, pp. 26-33, March/April 2022.
- [2] S. Aslam, M. P. Michaelides, and H. Herodotou, "Internet of Ships: A Survey on Architectures, Emerging Applications, and Challenges," IEEE IoT Journal, vol. 7, no. 10, pp. 9714-9727, May 2020.
- [3] 임민중, 김기현, 나영진, 소프트웨어 중심 선박을 위한 분산 데이터 수집, 한국통신학회 동계학술대회, 2025.