

정보 최신성 보장을 위한 지능형 네트워크 기술 동향

Intelligent Network Technical Trends for Ensuring Information Freshness

윤신웅 (S.W. Yun, syun4656@etri.re.kr)

이강찬 (K.C. Lee, chan@etri.re.kr)

이승윤 (S.Y. Lee, syl@etri.re.kr)

전략표준연구실 연구원

전략표준연구실 책임연구원/실장

표준연구본부 책임연구원/본부장

ABSTRACT

In next-generation communication networks, maintaining the freshness of information has emerged as a critical requirement, alongside traditional metrics such as latency and throughput. While conventional network designs focus on maximizing transmission efficiency or minimizing delay, they often overlook the timeliness of the received data. This report emphasizes the growing importance of information freshness, quantified by metrics such as Age of Information (AoI), and explores intelligent network architectures designed to optimize it. We present key freshness metrics, including average AoI, peak AoI, AoI violation probability, and Age of Incorrect Information, and examine their applicability across various network scenarios. The study also investigates reinforcement learning-based scheduling strategies that adapt to dynamic environments, with specific applications in cellular, sensor, and unmanned aerial vehicle networks. Our analysis demonstrates that intelligent algorithms, particularly those using deep reinforcement learning, are effective in managing resource constraints and optimizing update timing to preserve information freshness.

KEYWORDS Information freshness, intelligent network, reinforcement learning, 강화학습, 정보 최신성, 지능형 네트워크

I. 서론

차세대 네트워크 환경에서는 실시간성과 신뢰성이 요구되는 다양한 응용 분야의 등장에 따라, 정보

의 최신성에 대한 중요성이 점차 부각되고 있다[1]. 특히 자율주행 차량, 스마트 공장, 실시간 금융 거래 등에서는 최신 정보를 바탕으로 빠르고 정확한 의사결정을 내리는 시스템의 구축이 필수적이다.

* DOI: <https://doi.org/10.22648/ETRI.2025.J.400407>

* 이 연구는 과학기술정보통신부와 정보통신기획평가원(IITP)이 지원하는 인공지능 표준전문연구실 사업을 통해 발생한 결과임(No. 2021-0-02077).



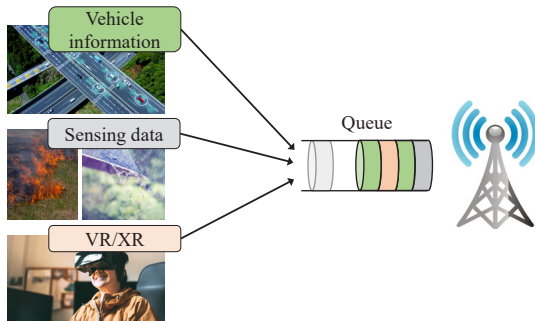
최근 인공지능 기술, 특히 강화학습(Reinforcement Learning)을 포함한 자율적인 의사결정 기법의 발전은 무선 네트워크 설계에 큰 변화를 일으키고 있다 [2,3]. 복잡하고 동적인 통신 환경에서 인공지능은 네트워크 상태의 변화를 실시간으로 학습하고 예측함으로써 사용자 요구에 최적화된 데이터 전송 전략을 도출할 수 있다. 이러한 기술은 기존 레이턴시(Latency)나 스루풋(Throughput) 향상을 넘어서, 네트워크가 전송하는 정보가 수신 시점에 얼마나 최신 상태인지를 보장하는 문제, 즉 정보 최신성이라는 새로운 과제를 해결하기 위한 핵심 도구로 주목받고 있다.

본고에서는 정보 최신성을 보장하기 위한 지능형 네트워크 설계 기술 동향을 다룬다. 먼저 정보 최신성의 필요성 및 개념을 소개 후, 이를 정량적으로 평가하기 위한 주요 지표들을 소개한다. 이후, 다양한 시나리오에서 정보 최신성 확보를 위한 지능형 네트워크 기술 사례를 소개한다.

II. 정보 최신성

1. 필요성

정보 최신성(Information Freshness)은 네트워크 시스템에서 데이터의 시간적 가치를 평가하는 개념



출처 게티이미지뱅크, 무단 전재 및 재배포 금지

그림 1 기지국에서의 정보 업데이트 시나리오

표 1 레이턴시 및 정보 최신성 비교

지연시간 (Latency)	<ul style="list-style-type: none"> 패킷이 송신단에서 수신단까지 도달하는 데 걸리는 시간 개별 정보 전송 건에 초점 End-to-End Delay, Round Trip Time 등이 지표로써 사용
정보 최신성 (Information Freshness)	<ul style="list-style-type: none"> 수신단이 보유한 정보가 얼마나 오래되었는지 나타냄 수신자 관점에서 정보의 최신 상태 유지 여부에 초점 정보의 나이(AoI) 등이 지표로써 사용

으로, 특히 실시간 통신과 IoT 환경 등에서 그 중요성이 대두되고 있다. 그림 1은 자율주행 차량 정보, 센서 데이터, VR 데이터 등 다양한 정보를 인근 기지국에 주기적으로 전송하는 예시를 보여준다. 이상적으로는 이러한 정보가 생성되는 즉시 수신자가 해당 정보를 얻을 수 있어야 하지만, 실제 네트워크 환경에서는 전송 지연이 존재하며, 이에 더해 큐(Queue)에 쌓인 다른 정보 패킷들의 영향을 받는다 [4]. 이 과정에서 하나의 센서가 특정 정보를 지나치게 자주 갱신하고 전송하면, 큐잉 레이턴시에 의해 다른 중요 정보의 전송이 지연되거나 누락될 수 있다.

따라서 기존 통신 지표로 널리 쓰이는 레이턴시 및 스루풋에 더해, 네트워크 자원을 효율적으로 활용하여 정보를 업데이트하는 최적의 전략이 요구된다. 표 1에서는 레이턴시 정보 최신성 비교를 통해 그 차이를 설명한다.

2. 정보의 나이

정보의 나이(AoI: Age of Information)는 정보 최신성을 측정하는 대표적인 지표로써 기존 레이턴시에 비해 데이터가 생성된 시점으로부터 현재까지의 경과 시간을 포괄적으로 고려한다[5]. 그림 2는 시간에 따른 정보의 나이 변화 예시를 보여준다. 송신단

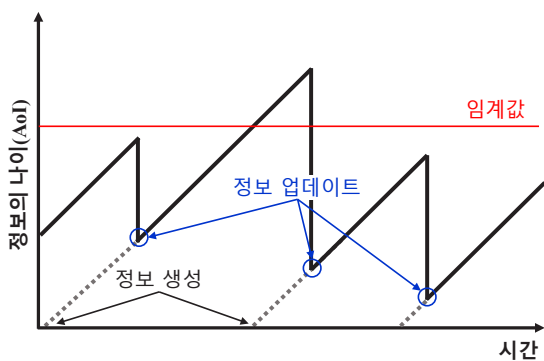


그림 2 시간에 따른 정보의 나이(AoI)

에서 정보가 생성된 직후로부터 시간에 따라 지속적으로 나이가 증가하며, 수신단에서 정보를 성공적으로 업데이트하는 순간 나이가 감소한다. 이러한 특징에 의해, 정보의 나이 그래프는 톱니 형태를 띠게 된다. 또한, 나이가 임계값을 초과하면 정보의 유효성이 급격히 떨어져 더 이상 신뢰할 수 없게 되므로, 최적의 업데이트 주기를 결정하도록 네트워크 설계가 요구된다.

3. 정보 최신성 지표

정보 최신성을 평가하기 위해 정보의 나이를 활용하여 상황과 목적에 따라 다양한 지표가 활용될 수 있다.

- **평균 AoI(Average AoI):** 일정 시간 구간 동안 평균 정보의 나이를 나타내는 지표이다. 해당 지표는 장기간 동안 시스템 전반적으로 최신성 성능을 요구하는 상황에 도입된다[6].
- **피크 AoI(Peak AoI):** 정보의 나이가 새롭게 갱신되기 직전까지 도달한 최댓값을 나타내는 지표이다. 해당 지표는 정보의 나이가 일정 임계값을 초과하지 않도록 보장해야 하는 상황에 도입된다[6].
- **AoI 위반 확률(AoI Violation Probability):** 전

체 작동 시간 중 정보의 나이가 임계값을 초과하는 시간의 비율을 나타내는 지표이다. 해당 지표는 피크 AoI와 유사하게 정보의 나이가 일정 임계값을 초과하지 않도록 보장해야 하는 상황에 도입된다[7].

- **불일치 정보의 나이(AoII: Age of Incorrect Information):** 정보가 실제 상태와 불일치한 상태로 유지된 기간을 나타내는 지표이다. 해당 지표는 기존 정보의 나이에 오차 함수를 추가한 형태로 표현되며, 정보의 최신성 및 정확성을 종합적으로 고려해야 하는 상황에 도입된다[8].
- **비선형 AoI(Non-Linear AoI):** 시간에 따라 정보의 가치가 선형적으로 변화하는 그림 2와 달리, 정보의 가치는 시간이 지날수록 급격히 감소하거나 비교적 천천히 감소할 수 있다. 비선형 AoI는 지수형(Exponential), 로그형(Logarithmic)과 같이 나이가 비선형적으로 증가하는 경우를 고려하여 정보의 가치를 더 정밀하게 모델링한다[9].

위 지표들은 각각 독립적으로 사용될 수도 있지만, 둘 이상을 함께 고려하여 시스템을 설계할 수도 있다. 예를 들어, 비선형 AoI 모델을 기반으로 평균 나이를 도출하거나, 평균 AoI와 피크 AoI가 각각 특정 값을 초과하지 않도록 만족시키는 스케줄링이 필요할 수 있다. 이처럼 지표 간 관계를 고려함으로써 정보의 최신성과 실용성을 더욱 정교하게 평가하고 반영할 수 있다.

III. 최신성 보장을 위한 강화학습 기반 의사결정

초창기 정보 최신성 보장 기술에서는 네트워크

트래픽 분포를 이용한 통계적 분석을 기반으로 최적의 정보 업데이트 주기를 도출하는 방식이 널리 활용되었다[4,5]. 그러나 실제 네트워크 환경은 시간이 지남에 따라 유동적으로 변화하며, 패킷 도착률, 채널 상태 등 다양한 요소가 동적으로 상호작용한다. 기존 통계 분석 기술은 이러한 변화에 즉각적으로 대응하지 못하며, 이는 실시간 응용 환경에서 성능 저하를 초래할 수 있다.

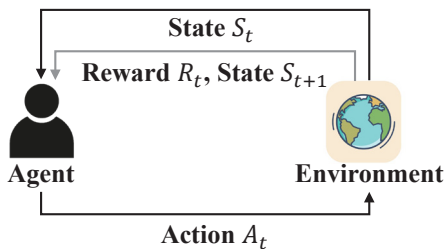
이러한 배경 속에서, 최근 강화학습 기반의 동적 스케줄링 기법이 정보 최신성 보장을 위한 지능형 알고리즘으로 주목받고 있다. 강화학습은 환경과의 상호작용을 통해 획득하는 보상을 기반으로 정책을 학습하므로, 시스템 상태가 변하는 경우에도 적응적으로 업데이트 전략을 조절할 수 있다[10].

본 장에서는 강화학습 기반 스케줄링을 모델링하는 마르코프 결정 과정(MDP: Markov Decision Process) 및 지능형 네트워크 설계에 주로 사용되는 강화학습 알고리즘을 소개한다.

1. 마르코프 결정 과정

마르코프 결정 과정은 결과가 확률적으로 결정되는 의사결정 문제를 모델링하는 데 사용되는 프레임워크이다[11]. 해당 과정을 통해 강화학습에서 시간별 상호작용을 표현할 수 있다.

그림 3은 주어진 시점에서 강화학습 상호작용을



출처: 게티이미지뱅크, 무단 전재 및 재배포 금지

그림 3 에이전트-환경 간 상호작용

보여준다. 에이전트(Agent)는 학습을 수행하는 주체로, 주어진 환경(Environment)에서 최적의 행동을 선택하기 위해 정책을 학습한다. 반면, 환경은 에이전트의 행동에 따라 그 상태가 변화하고, 이에 대한 보상을 에이전트에게 반환한다. 이러한 상호작용은 다음 요소들을 중심으로 구성된다.

- **에이전트(Agent):** 현재 상태를 기반으로 액션을 결정하고 학습을 수행하는 주체이다.
- **상태(State):** 에이전트는 환경의 현재 상태를 관찰한다. 예를 들어, 네트워크 스케줄링 문제에서는 현재 정보의 나이와 같은 실시간 정보가 포함될 수 있다.
- **액션(Action):** 에이전트는 현재 상태를 기반으로 액션을 선택하고, 환경을 변화시킨다.
- **보상(Reward):** 에이전트가 액션을 수행하면, 환경은 그 결과에 따라 보상을 반환한다. 이는 일반적으로 시스템의 성능을 정량화한 값으로 정해진다. 예를 들어, 여러 지표의 합이나 평균을 기반으로 보상을 설정할 수 있다. 또한, 특정 조건이 충족되지 않거나 기준 이하일 경우, 페널티를 부여하거나 보상을 감소하는 방식으로 네트워크를 디자인할 수 있다.
- **정책(Policy):** 현재 상태가 주어졌을 때, 에이전트가 어떤 액션을 선택할지 결정하는 규칙을 의미한다.
- **상태 전이 확률:** 에이전트의 액션은 다음 상태를 확률적으로 결정한다. 즉, 정책에 따라 액션을 취했을 때 다음 상태가 결정되는 확률을 상태 전이 확률이라 하며, 이는 환경의 동적 특성을 반영한다.

2. 강화학습 알고리즘

강화학습의 목적은 에이전트가 현재 상태가 주어

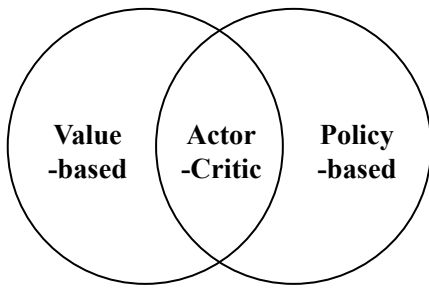


그림 4 강화학습 분류

졌을 때 어떤 액션을 취할지 그 정책을 학습하여 누적되는 보상을 최대화하는 것이다. 이때, 누적 보상은 시간이 지남에 따라 미래 보상의 가치가 점차 감소하도록 설계된 감쇠 보상의 합으로 정의된다. 이를 위해 도입되는 요소가 감쇠 인자(Discount Factor)이며, 0에서 1 사이의 값을 가진다. 감쇠 인자는 미래 보상이 현재 결정에 미치는 영향을 조절한다. 값이 1에 가까울수록 먼 미래의 보상까지 중요하게 고려하는 장기적 학습을 유도하고, 0에 가까울수록 즉각적 보상에 초점을 맞춘 단기적인 학습을 유도한다. 강화학습에서는 에이전트가 상태-액션 경로를 따라가며 얻는 감쇠 보상 합계의 기댓값을 최대화하는 것이 핵심 목표가 된다.

이를 위해 에이전트는 주기적으로 정책을 업데이트함으로써 더 좋은 보상을 얻을 수 있도록 학습한다. Model-Free 강화학습 방식은 그림 4와 같이 크게 세 가지로 분류된다[12].

2.1 가치 기반 강화학습

가치 기반(Value-Based) 강화학습은 가치함수를 도입함으로써 정책을 간접적으로 학습한다. 현재 상태 및 액션에 따라 기대되는 보상을 가치함수로 두고, 에이전트는 주어진 상태에서 가장 높은 가치함수를 얻을 수 있는 액션을 선택하는 방식으로 정책을 업데이트한다.

해당 방식은 상대적으로 구현이 간단하고 이산 액션 공간을 가지는 문제에서 좋은 성능을 보이지만, 연속적이거나 고차원의 액션을 가지는 환경에 적용하기 어렵다. 대표적인 가치 기반 강화학습 알고리즘으로는 Q-Learning[13], Deep Q-Network(DQN)[14] 등이 있다.

2.2 정책 기반 강화학습

정책 기반(Policy-Based) 강화학습은 가치함수를 추정하는 대신 정책 자체를 직접 학습하는 방식이다. 정책이 신경망과 같은 함수 근사기로 파라미터화되며, 주어진 상태에서 어떤 행동을 선택할지 출력하는 구조를 가진다. 에이전트는 정책의 파라미터를 조정함으로써 기대 누적 보상을 최대화하도록 학습한다.

정책 기반 강화학습은 연속적이거나 고차원적인 액션 공간을 다루는 환경에서 강점을 보인다. 또한, 확률 정책을 학습할 수 있으므로 불확실한 상황에서의 탐색 능력이 뛰어나며, 정책을 명시적으로 표현하므로 다른 전략과의 결합이 유연하다. 그러나, 샘플 비효율성으로 인해 많은 데이터를 필요로 하고, 정책 그래디언트 추정은 높은 분산으로 인해 학습이 불안정해질 수 있다. 대표적인 정책 기반 강화학습 알고리즘으로는 REINFORCE[15], Proximal Policy Optimization(PPO)[16] 등이 있다.

2.3 Actor-Critic

Actor-Critic 강화학습은 가치 기반 및 정책 기반 강화학습을 결합한 접근법이다. 해당 방식은 정책을 직접 학습하는 Actor와 그 정책의 성능을 평가하는 Critic이 함께 작동한다. Actor는 주어진 상태에서 어떤 액션을 선택할지 결정하며, Critic은 Actor가 내린 액션의 가치를 추정하여 정책을 개선할 수 있도록 피드백을 제공한다. 이 과정에서 Critic은 보상을

바탕으로 가치함수를 업데이트하고, Actor는 그 정보를 활용해 정책 파라미터를 조정한다.

해당 방식은 앞의 두 가지 방식을 결합하여 보다 안정적이며, 복잡한 환경에서도 높은 성능을 얻을 수 있다. 대표적인 Actor-Critic 기반 강화학습 알고리즘으로는 Deep Deterministic Policy Gradient (DDPG)[17], Soft Actor-Critic(SAC)[18]이 있다.

IV. 최신성 보장을 위한 지능형 네트워크 설계

정보 최신성을 보장하기 위한 네트워크 설계는 최근 셀룰러, 센서, 비지상 네트워크 등 다양한 응용 환경을 중심으로 활발히 연구되고 있다. 이러한 시스템에서는 실시간 정보 전달의 신뢰성과 정확성이 핵심 과제로 대두되며, 최신성 보장 없이는 안정적인 서비스 품질을 확보하기 어렵다. 그러나 기존의 네트워크 설계 방식은 정적인 정책이나 통계 기반의 전송 주기 설정에 의존하고 있어, 트래픽 변동이나 네트워크 상태 변화에 실시간으로 대응하는 데 한계가 있다. 특히 정보가 지나치게 자주 또는 불필요하게 전송될 경우 오히려 자원 낭비와 레이턴시 증가로 이어져, 다른 중요한 정보의 신속한 전달을 방해할 수 있다.

이러한 한계를 극복하고, 변화하는 환경 속에서도 적응적으로 정보의 최신성을 유지하기 위해 최근에는 강화학습 기반의 네트워크 디자인이 도입되는 추세이다. 본 장에서는 강화학습 기반의 최신성 보장 네트워크 설계 방안을 중심으로, 마르코프 결정 과정을 활용한 문제 정의 방식과 대표적인 학습 알고리즘들의 구조 및 특징을 소개한다. 이를 위해 셀룰러, 센서, 무인기 네트워크에서의 최신성 보장을 위한 지능형 네트워크 설계 예시를 보여준다.

1. 셀룰러 네트워크

셀룰러 네트워크는 대규모 사용자 단말이 기지국과 연결되어 지속적으로 데이터를 송수신하는 구조로, 실시간 모니터링, 스마트 팩토리, 차량 통신 등 다양한 실시간 응용 환경에 활용된다. 이러한 환경에서는 단말이 전송하는 상태 정보가 목적지에 얼마나 신속히 도달하는지가 시스템의 반응성과 안정성을 결정짓기 때문에, 정보의 최신성 유지가 중요한 과제가 된다. 특히 5G 환경에서는 네트워크가 초저지연 및 고신뢰성 통신을 지원해야 하므로, 주어진 무선 자원을 효율적으로 활용하여 각 단말의 전송을 조정하는 지능형 스케줄링이 요구된다.

셀룰러 네트워크의 정보 최신성 최적화에는 다양한 제약이 존재한다. 예를 들어, 일부 단말은 업링크 전송을 위한 무선 자원이 제한적으로 배정되기도 하며, 서로 다른 통신 경로를 동시에 활용할 수도 있다. 이에 따라, 각 단말의 채널 상태, 큐 길이 등 다양한 시스템 상태 정보를 고려해 전송 경로를 선택하거나 업데이트 요청 여부를 결정해야 한다.

표 2는 정보 최신성 보장을 위한 셀룰러 네트워크에서의 마르코프 결정 과정을 보여준다[19-21]. 주로 기지국이 에이전트로서 결정을 내리나, 업링크 환경에서는 유저단에서 에이전트 역할을 수행하기도 한다. 에이전트는 유저측에서의 정보의 나이 및

표 2 최신성 보장을 위한 셀룰러 네트워크 마르코프 결정 과정

에이전트	<ul style="list-style-type: none"> • 기지국 • 유저
상태	<ul style="list-style-type: none"> • 정보의 나이 • 버퍼 길이 • 과거 스루풋 히스토리
액션	<ul style="list-style-type: none"> • 대역폭 할당 • 리소스 스케줄링
보상 구성 요소	<ul style="list-style-type: none"> • 정보 최신성 지표 • 스루풋

버퍼 길이, 과거 스루풋 등 정보를 기반으로 대역폭 할당, 리소스 스케줄링을 수행한다. 이에 따라 각 유저단에서의 정보의 나이 및 스루풋 등을 결합하여 보상을 획득, 이를 최대화하는 방향으로 학습을 진행한다. 심층(Deep) 강화학습을 이용하는 경우, 현재 상태를 입력받아 인공 신경망을 통해 적합한 액션을 선택하고, 이를 환경에 반영 후 보상을 얻어 학습에 이용한다.

2. 센서 네트워크

센서 네트워크에서는 여러 센서가 분산된 위치에 서 실시간으로 데이터를 수집하고, 이를 중앙 노드로 전송함으로써 전체 시스템의 상태 정보를 구성한다. 각 센서는 개별적으로 정보를 생성하고 이를 전송하는 역할을 수행하며, 중앙 노드는 네트워크 내 각 센서의 정보 상태를 추적하고 정보의 나이를 관리한다. 해당 네트워크에서는 각 센서로부터 얼마나 자주, 그리고 얼마나 정확한 시점의 데이터가 도착하는지가 전체 네트워크의 정보 최신성 수준을 결정하게 된다.

그러나 센서는 배터리 용량이 제한적이며, 데이터 수집 및 통신 과정에서 지속적으로 에너지를 소모한다. 따라서 각 센서는 자주 데이터를 생성하고 전송하는 대신, 자신의 현재 배터리 및 최신성 상태를 고려하여 전략적으로 동작을 선택해야 한다.

정보 최신성 보장을 위한 센서 네트워크 마르코프 결정 과정은 주로 표 3과 같이 이루어진다[7,22,23]. 중앙 노드가 에이전트로서 센서들의 액션을 결정하거나, 각 센서가 에이전트로서 독자적으로 액션을 결정할 수 있다. 센서는 주로 낮은 연산 능력을 가지고 있으므로, 모델 학습은 중앙 노드 측에서 수행 후 학습된 모델을 각 센서에 배포한다. 각 센서는 현재 정보의 나이, 배터리 상태 등 실시간 정

표 3 최신성 보장을 위한 센서 네트워크 마르코프 결정 과정

에이전트	<ul style="list-style-type: none"> • 센서 • 중앙 노드
상태	<ul style="list-style-type: none"> • 센서 정보의 나이 • 센서 배터리 상태 • 추정 무선 채널 상태
액션	<ul style="list-style-type: none"> • 센서 정보 업데이트 여부 결정
보상 구성 요소	<ul style="list-style-type: none"> • 정보 최신성 지표 • 센서 에너지 소비량

보를 기반으로 센싱 및 정보 업데이트 여부를 결정 후, 센서들의 정보의 나이에 따라 보상을 받게 된다. 이는 단일 센서 수준의 최적화가 아닌, 네트워크 전체의 협조적인 최신성 유지와 연계되어 실질적인 성능 향상을 가능케 한다.

3. 무인기 네트워크

무인기는 기동성과 높은 가시성(LoS: Line-of-Sight)을 바탕으로, 실시간 감시, 재난 대응, 환경 모니터링 등 다양한 임무에 활용되고 있다. 해당 통신 환경에서는 무인기가 공중에서 데이터를 수집하고 이를 지상 수신 노드로 전달함으로써 전체 네트워크의 정보 흐름을 형성한다. 특히 무인기는 자율 비행 경로 결정, 에너지 소모, 그리고 동적 채널 상태와 같은 다양한 변수를 동시에 고려해야 하므로, 정보 최신성을 유지하기 위한 설계가 복잡하다.

무인기는 배터리 용량과 비행 시간에 제약이 있으며, 고도 및 위치 변화에 따라 통신 성공률이 크게 달라진다. 또한, 무인기 간 충돌 회피, 데이터 수집 범위 최적화, LoS 확률 제어 등의 요소도 실시간으로 반영되어야 한다.

표 4는 정보 최신성 보장을 위한 무인기 네트워크 마르코프 결정 과정을 보여준다[24,25]. 각 무인기가 에이전트로서 액션을 결정하거나, 중앙 노드에 서 다수 무인기의 액션을 조종할 수 있다. 위치가 고

표 4 최신성 보장을 위한 무인기 네트워크 마르코프 결정 과정

에이전트	<ul style="list-style-type: none"> • 무인기 • 중앙 노드
상태	<ul style="list-style-type: none"> • 무인기/지상 노드 정보의 나이 • 무인기 위치 • 무인기 배터리
액션	<ul style="list-style-type: none"> • 무인기 경로(이동 방향 및 거리) • 업데이트 지상 노드 결정
보상 구성 요소	<ul style="list-style-type: none"> • 정보 최신성 지표 • 무인기 에너지 소비량 • 무인기 충돌/경로 이탈 페널티

정되어 있는 센서와 달리, 무인기는 시간에 따라 경로를 이동하므로 정보의 나이, 배터리 상태에 더해 현재 위치를 상태로 이용한다. 해당 정보를 바탕으로 이동 방향 및 거리를 포함한 무인기 경로를 지정하고, 업데이트가 필요한 지상 노드를 액션으로써 결정한다. 정보 최신성 지표, 무인기 에너지 소비량에 따라 보상이 결정되며 무인기 간 충돌 및 경로 이탈을 방지하기 위해 해당 상황 발생 시 큰 페널티를 주는 방향으로 보상을 설계하기도 한다.

V. 결론

본고에서는 정보 최신성의 개념과 이를 정량화하는 다양한 지표를 소개하고, 이를 효과적으로 보장하기 위한 지능형 네트워크 설계 방안을 다양한 환경에 걸쳐 분석하였다.

차량, 무인기, 웨어러블 디바이스 등 수많은 통신 단말이 연결되는 차세대 네트워크에서는 실시간으로 생성 및 전달되는 데이터의 양이 기하급수적으로 증가하는 추세이다. 이에 따라, 다수 기기의 정보 업데이트를 스케줄링함으로써 해당 단말들의 레이턴시, 신뢰도, 스루풋 등과 함께 최신성을 보장해 주어야 한다. 강화학습을 필두로 한 지능형 알고리즘은 실시간으로 변하는 환경에 적응하며 자율적인 의사결정을 바탕으로 종합적인 시스템 성능을 보장하는 열쇠가 될 것이다. 이는 본고에서 소개한 네트워크 예시 이외에도 위성 통신, 차량 통신 등 다양한 네트워크 환경에 활용될 전망이다.

용어해설

Aol 정보 최신성을 정량적으로 평가하는 핵심 지표로써 정보가 생성된 이후 수신단에서 업데이트하기까지 걸리는 시간을 나타냄
MDP 순차적인 의사결정 문제를 풀기 위해 강화학습에 사용되는 수학적 모델

참고문헌

- [1] J. Chu, "Keeping data fresh for wireless networks," MIT News, 2018. 6. 4. <https://news.mit.edu/2018/keeping-data-fresh-wireless-networks-0605>
- [2] Y. Chen et al., "Reinforcement learning meets wireless networks: A layering perspective," IEEE Internet Things J., vol. 8, no. 1, 2021, pp. 85-111.
- [3] NVIDIA, "NVIDIA and Telecom Industry Leaders to Develop AI-Native Wireless Networks for 6G," NVIDIA Newsroom, 2025. 3. 18. <https://nvidianews.nvidia.com/news/nvidia-and-telecom-industry-leaders-to-develop-ai-native-wireless-networks-for-6g>
- [4] I. Kadota et al., "Optimizing age of information in wireless networks with throughput constraints," in Proc. IEEE Conf. Comput. Commun. (INFOCOM), (Honolulu, HI, USA), Apr. 2018, pp. 1-9.
- [5] S. Kaul et al., "Real-time status: How often should one update?," in Proc. IEEE Conf. Comput. Commun. (INFOCOM), (Orlando, FL, USA), Mar. 2012, pp. 2731-2735.
- [6] Q. Abbas et al., "A comprehensive survey on age of information in massive IoT networks," Comput. Commun., vol. 197, 2023, pp. 199-213.
- [7] S.W. Yun et al., "Learning-based sensing and computing decision for data freshness in edge computing-enabled networks," IEEE Trans. Wirel. Commun., vol. 23, no. 9, 2024, pp. 11386-11400.
- [8] A. Maatouk et al., "The age of incorrect information: A new performance metric for status updates," IEEE/ACM Trans. Netw., vol. 28, no. 5, 2020, pp. 2215-2228.
- [9] A. Kosta et al., "Age and value information: non-linear age case," in Proc. IEEE Int. Symp. Inform. Theory (ISIT), (Aachen, Germany), Jun. 2017, pp. 1-5.
- [10] J. Murel and E. Kavlakoglu, "What is reinforcement learning?," IBM, 2024. 3. 25. <https://www.ibm.com/think/topics/reinforcement-learning>
- [11] R. Jagtap, "Understanding the Markov Decision Process (MDP)," builtin, 2024. 8. 13. <https://builtin.com/machine-learning/markov-decision-process>
- [12] Papers in 100 Lines of Code, "Value-Based vs Policy-Based Reinforcement Learning," Medium, 2024. 11. 21. <https://papers-100-lines.medium.com/value-based-vs-policy-based-reinforcement-learning-92da766696fd>
- [13] C.J.C.H. Watkins and P. Dayan, "Q-Learning," Mach. Learn., vol. 8, 1992, pp. 279-292.
- [14] V. Mnih et al., "Human-level control through deep reinforcement learning," Nature, vol. 518, 2015, pp. 529-533.
- [15] R.J. Williams, "Simple statistical gradient-following algorithms for connectionist reinforcement learning," Mach. Learn., vol. 1992, pp. 229-256.
- [16] J. Schulman et al., "Proximal policy optimization algorithms," arXiv preprint, 2017. doi: 10.48550/arXiv.1707.06347
- [17] T.P. Lillicrap et al., "Continuous control with deep reinforcement learning," arXiv preprint, 2015. doi: 10.48550/arXiv.1509.02971
- [18] T. Haarnoja et al., "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in Proc. Int. Conf. Mach. Learn. (ICML), (Stockholm, Sweden), Jul. 2018, pp. 1861-1870.
- [19] C.C. Wu et al., "Design of AoI-aware 5G uplink scheduler using reinforcement learning," in Proc. IEEE 5G World Forum (5GWF), Oct. 2021, pp. 176-181.
- [20] A.I. Rasul and H. Beitollahi, "Optimizing information freshness in dual-connection cellular networks for real-time applications," IEEE Access, vol. 12, 2024, pp. 94164-94175.
- [21] Z. Mlika and S. Cherkaoui, "Deep deterministic policy gradient to minimize the age of information in cellular V2X communications," IEEE Trans. Intell. Transp. Syst., vol. 23, no. 12, 2022, pp. 23597-23612.
- [22] X. Wu et al., "Deep reinforcement learning for IoT networks: Age of information and energy cost tradeoff," in Proc. IEEE Global Commun. Conf. (GLOBECOM), (Taipei, Taiwan), Jan. 2021, pp. 1-6.
- [23] T.W. Song and Y.W. Kyung, "Deep-reinforcement-learning-based age-of-information-aware low-power active queue management for IoT sensor networks," IEEE Internet Things J., vol. 11, no. 9, 2024, pp. 16700-16709.
- [24] M. Yi et al., "Deep reinforcement learning for fresh data collection in UAV-assisted IoT networks," in Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS), (Toronto, ON, Canada), Jul. 2020, pp. 716-721.
- [25] M. Samir et al., "Age of information aware trajectory planning of UAVs in intelligent transportation systems: A deep learning approach," IEEE Trans. Veh. Technol., vol. 69, no. 11, 2020, pp. 12382-12395.