

토큰 샘플링 기반 동적 프레임 슬롯 알고하

조영제, 황경호*

국립한밭대학교 컴퓨터공학과

yeongjejo@edu.hanbat.ac.kr, *gabriel@hanbat.ac.kr

Dynamic Framed Slotted ALOHA based on Thompson Sampling

Yeong-Je Jo, Gyung-Ho Hwang*

Hanbat National University, Dept. Computer Engineering

요약

본 논문에서는 Q-learning 기반의 Slotted ALOHA 프로토콜인 ALOHA-Q 방식에서 노드의 수와 프레임 당 슬롯의 수가 차이가 있을 때 throughput 성능이 낮아지는 점을 보완하기 위하여 Slotted ALOHA에 토큰 샘플링 강화학습 기법을 적용하여 샘플링 값에 따라 프레임 내의 슬롯 수를 조절하는 방안을 제안하였다. 슬롯에서의 패킷 충돌 확률과 빈 슬롯 발생 확률을 모두 감소시켜 Slotted ALOHA, ALOHA-Q 방식 대비 높은 throughput을 보여주는 것을 시뮬레이션을 통한 성능 분석으로 확인하였다.

I. 서론

Slotted ALOHA에 Q-learning 기반 강화학습을 적용한 ALOHA-Q 방식은 각각의 노드들이 프레임 내의 전송 성공률이 높은 슬롯에 패킷을 전송하게 하여 Slotted ALOHA 대비 성능을 높였지만 프레임 내의 슬롯의 개수가 노드 수보다 많으면 빈 슬롯이 생기고, 슬롯의 개수가 노드 수보다 적으면 슬롯에서 패킷 충돌 확률이 높아져 throughput 성능이 저하되는 문제가 있다.

본 논문에서는 Slotted ALOHA에 토큰 샘플링 기반의 강화학습을 적용하여 노드들이 패킷 전송 성공 확률이 높은 슬롯에 패킷을 전송할 수 있도록 학습할 뿐만 아니라, 노드의 수에 따라 프레임 내의 슬롯 수를 조절하여 빈 슬롯을 줄이고 패킷 충돌 확률을 감소시켜 성능을 높이는 방식을 제안한다.

II. 본론

2.1 Dynamic Framed Slotted ALOHA (DFSA)

Framed Slotted ALOHA에서는 프레임을 여러 개의 슬롯으로 나누어 슬롯의 시작점에서만 패킷을 전송할 수 있게 하여 Pure ALOHA 대비 성능을 높일 수 있지만, 프레임당 슬롯의 수를 고정으로 할당하여 슬롯 개수 대비 노드 개수가 많거나 작아질수록 성능이 떨어졌다. DFSA는 이러한 문제를 보완하여 프레임 내의 슬롯 수를 조절해가며 성능을 높이는 방식이다. Vogt 알고리즘은 프레임 내에서 충돌이 발생한 슬롯 수를 이용하여 다음 프레임 크기를 조절하는 알고리즘으로 슬롯에 충돌이 발생하면 해당 슬롯에는 최소한 2개 이상의 노드가 전송을 시도한 것을 의미하므로 충돌이 발생한 슬롯 수에 2를 곱한 만큼 다음 프레임 크기를 증가시킨다.[1]

DFSA 방식에서는 프레임의 크기 조절은 가능하지만, 노드가 랜덤하게 슬롯을 선택하여 성능의 한계가 존재한다.

2.2 ALOHA-Q

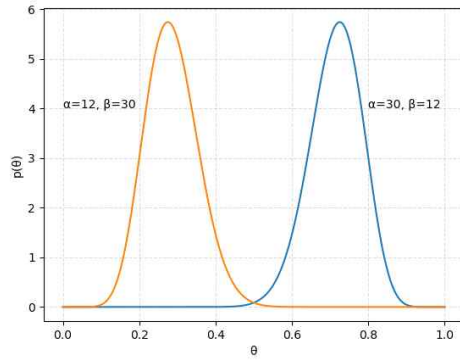
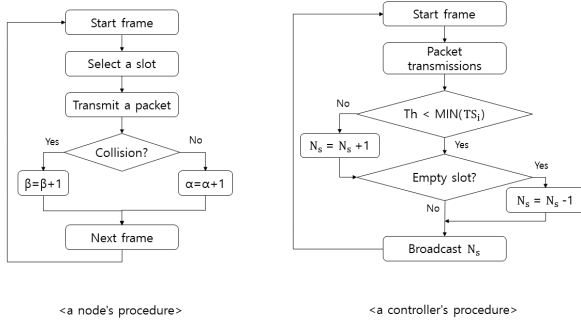
ALOHA-Q 방식은 Slotted ALOHA에 Q-learning 기반 강화학습을 적용하여 이전 프레임에서 선택한 슬롯에 대해 패킷 전송의 보상 값에 따라 노드들이 프레임 내의 슬롯 중 전송 성공률이 가장 높은 슬롯을 학습하는 방식이다. 이상적인 경우 다수의 노드가 서로 다른 슬롯에 패킷을 전송하여 패킷 충돌이나 빈 슬롯이 발생할 확률을 낮추어 성능을 높인다. 하지만 프레임 내의 슬롯 수가 고정이므로 노드 수와 차이가 날수록 성능이 낮아진다.[2][3]

2.3 토큰 샘플링 강화학습 기반의 DFSA

토큰 샘플링은 (1)과 같은 베타 분포를 이용한 알고리즘으로 α 값과 β 값을 이용하여 action에 대한 확률분포를 만들어 무작위로 값을 샘플링하여 그중 가장 큰 값을 가지는 action을 정한다. $\Gamma()$ 는 감마 함수를 의미하며 θ 는 action에 대한 보상을 받을 확률을 나타내며 $1-\theta$ 는 보상을 받지 못할 확률을 나타낸 값이다. Action에 대한 보상을 받게 되면 α 값이 증가하여 분포가 1에 근접하게 되어 샘플링 값이 높을 확률이 높아지고 보상을 받지 못하면 β 값이 증가하면서 분포가 0에 근접하게 되어 샘플링 값이 낮을 확률이 높아지게 된다.[4]

$$p(\theta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} \quad (1)$$

본 논문에서는 Slotted ALOHA에 토큰 샘플링을 적용한 DFSA 방식을 제안한다. 노드들은 각자 프레임 내의 모든 슬롯에 대해 베타 분포를 가지고 α 는 해당 슬롯에 패킷 전송을 성공한 횟수를 나타내고 β 는 패킷 전송을 실패한 횟수를 나타낸다. <그림 1>은 두 슬롯에 대한 노드의 베타 분포를 표현한 예시다. 패킷 전송에 성공하면 α 값이 증가하여 해당 슬롯에 대한 베타분포의 확률밀도함수의 값이 θ 가 1에 가까워질 때 높아지고,

그림 1. α , β 값에 따른 베타분포그림 2. 제안하는 톰슨 샘플링 강화학습 기반
Dynamic Framed Slotted ALOHA 동작

패킷 전송에 실패하면 β 값이 증가하여 θ 가 0에 가까울 때 확률밀도함수가 높아진다. 슬롯에 패킷을 전송하지 않으면 분포에는 변화가 생기지 않는다. 프레임이 진행될수록 베타 분포의 확률밀도함수의 폭이 좁아져 전송 성공확률이 높은 슬롯의 샘플링 값은 1에 가까워지고, 나머지 슬롯에 대한 샘플링 값은 0에 가까운 값이 되어 전송 성공 확률이 높은 슬롯을 선택할 가능성이 높아진다.

<그림 2>는 제안한 방안의 한 프레임 동안의 동작을 나타낸다. 모든 노드는 프레임 시작 시 톰슨 샘플링 알고리즘을 이용하여 각 슬롯의 샘플링 값을 구하고 그 중 가장 큰 샘플링 값의 슬롯을 선택하여 패킷을 전송하고 성공 여부에 따라 α , β 값을 업데이트한다. 프레임의 크기를 조절하는 컨트롤러는 노드들이 선택한 슬롯에 대한 샘플링값(TS_i , 노드 i 가 선택한 슬롯의 톰슨 샘플링 값)들의 최솟값이 미리 설정한 임계값 Th 보다 낮으면 전송 성공률이 높은 슬롯을 선택할 수 없을 확률이 높으므로 프레임의 슬롯 수를 의미하는 N_s 를 1 증가시킨다. 반대로 노드들의 샘플링 값이 모두 Th 보다 높지만 빈 슬롯이 발생하면 불필요한 슬롯이 존재할 확률이 높으므로 N_s 를 1 감소시킨다.

2.4. 시뮬레이션을 통한 성능 분석

본 논문에서 제안하는 방식의 성능을 분석하기 위해 Python을 이용하여 시뮬레이션을 진행하였고 임계값 Th 는 0.5로 설정하였다.

<그림 3>은 노드 수에 따른 에피소드별 슬롯 수를 나타낸 것이다. 에피소드는 한 프레임 동안의 접속을 의미하며 첫 에피소드에서 프레임의 크기는 10개 슬롯으로 설정하였다. 에피소드가 증가함에 따라 노드 개수에 맞게 프레임의 슬롯 수가 조절되는 것을 볼 수 있다.

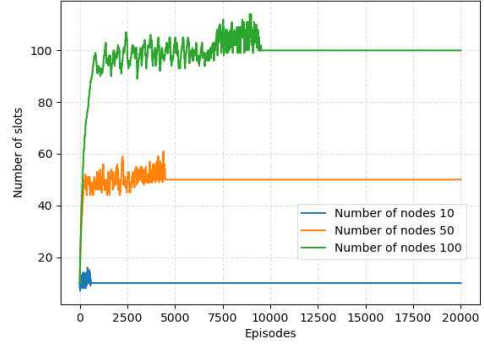


그림 3. 에피소드별 한 프레임의 슬롯 수

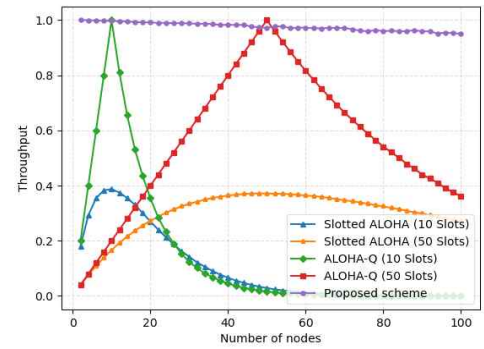


그림 4. 노드 수에 따른 throughput 비교

<그림 4>는 본 논문에서 제안하는 방식과 다른 방식의 성능을 비교한 것으로 100,000 에피소드 동안의 평균 throughput을 나타낸다. Throughput은 시뮬레이션 동안의 모든 슬롯 개수 대비 패킷 전송을 성공한 슬롯 개수의 비율이다. Slotted ALOHA, ALOHA-Q 방식 모두 프레임당 슬롯의 개수를 고정해 슬롯 수와 노드 수의 차이가 클수록 throughput이 감소하는 걸 확인하였고 본 논문에서 제안한 방식은 톰슨 샘플링 강화학습을 사용하여 프레임의 슬롯 개수를 동적으로 변화시키는 과정에서 노드의 수가 증가할수록 최적의 슬롯 개수를 찾는데 더 많은 에피소드가 필요하여 throughput이 점차 감소하지만 모든 경우에서 throughput이 1.0에 가까운 성능을 보여주었다.

III. 결론

본 논문에서는 Slotted ALOHA에 톰슨 샘플링 강화학습을 적용한 DFSA 방식을 제안하였다. 톰슨 샘플링 값에 따른 프레임의 슬롯 개수를 동적으로 조절하고, 강화학습을 통해 전송 성공확률이 높은 슬롯을 결정하여 Slotted ALOHA와 ALOHA-Q 방식 대비 throughput 성능이 향상하였다.

참고 문헌

- [1] H. Vogt, "Efficient object identification with passive RFID tags," in Proc. Int. Conf. Pervasive Computing, pp. 98–113, 2002.
- [2] R. S. Sutton and A. G. Barto, *Reinforcement Learning An Introduction*, MIT press, 2018.
- [3] 이창규, 이승형, "강화학습 기반 Slotted ALOHA의 성능개선", 한국통신학회논문지, vol.45, no.11, pp.1886–1892, 2020.
- [4] D. J. Russo, B. V. Roy, A. Kazerouni, I. Osband and Z. Wen, "A Tutorial on Thompson Sampling," Foundations and Trends in Machine Learning, vol.11, no.1, pp.1–96, 2018.