

강화학습을 활용한 RFID 충돌 회피 알고리즘의 성능 향상

김 태 욱*, 황 경 호^o

Performance Enhancement of RFID Anti-Collision Algorithm Based on Reinforcement Learning

Tae-Wook Kim*, Gyung-Ho Hwang^o

요 약

RFID(Radio Frequency Identification) 시스템에서 태그가 자신의 ID를 리더기에 보낼 때, 전송되는 패킷들의 충돌을 줄이기 위해 한 프레임에 속한 슬롯의 개수를 변화시킬 수 있다. 태그 개수 대비 슬롯의 개수가 작거나 크게 되면 처리율이 감소하거나 모든 태그의 ID를 전송하는데 걸리는 시간이 길어진다. 본 논문에서는 RFID 시스템에서 프레임 크기의 변화를 강화학습 중 Q-Learning 기법을 통해 결정하는 방법을 사용하여 기존 방법보다 처리율과 태그 ID 전송 시간 관점에서 더 좋은 성능을 보여주었다.

Key Words : RFID, Reinforcement Learning, Q-Learning, Anti-collision algorithm

ABSTRACT

In the RFID (Radio Frequency Identification) systems, when a number of tags send their own IDs, the number of slots in a frame can be adjusted to alleviate the collisions. However, non-optimal frame size causes throughput decrease and significantly increases the time to complete transmissions. In this letter, we adopted the Q-Learning, one of reinforcement learning methods and the proposed method showed better performance in views of

throughput and transmission completion time compared previous studies.

I. 서 론

RFID 시스템은 공유 무선 통신 채널을 사용하는 무선 통신 방식 중 하나로 여러 개의 태그와 하나 이상의 리더(reader)로 구성된다. 태그는 리더의 신호에서 후방 산란 전자기 에너지를 사용하여 ID 정보를 전달하며, 동시에 두 개 이상의 태그가 ID를 전송하면 충돌이 발생한다. RFID 시스템의 충돌 회피 프로토콜은 트리 기반과 ALOHA 기반의 두 가지 방식이 있으며 EPC-C1G2 표준은 ALOHA 기반의 알고리즘을 사용한다. 본 논문에서는 ALOHA 기반의 RFID 충돌 회피 알고리즘을 사용할 때, 강화학습의 Q-learning 기법을 활용하여 프레임의 슬롯 개수를 조절하는 알고리즘을 제안한다.

II. 관련 연구

2.1 태그 개수 추정 알고리즘

RFID의 충돌 회피 방식들 중 ALOHA를 기반으로 하는 알고리즘에서 중요한 요소는 남아 있는 태그의 개수를 정확히 추정하는 것이다. [1]에서는 체비셰프 부등식에 근거한 기댓값 대신 측정값을 적용한 변형식을 사용한다. 한 프레임에서 충돌이 발생한 슬롯에 2.3922배를 한 값을 태그의 개수로 추정하고 있다. [2]에서는 태그 전송의 성공과 충돌, 유효시간을 고려한 처리율의 공식을 사용해서 한 프레임에서 충돌이 발생한 슬롯에 1.46배를 한 값을 태그의 개수로 추정하고 있다. 하지만, 정확한 태그 수 추정은 불가능하고, 이를 통한 슬롯 개수 조절이 최고의 성능을 보장하지 못한다.

2.2 EPC-C1G2 표준의 Q-Algorithm

EPC-C1G2(EPC-Global Class 1 Generation 2)는 RFID 장치를 상호 운용할 수 있는 RFID 무선 인터페이스 표준 프로토콜이다.^[3] 태그들로부터 수신한 ID의 충돌 여부를 매 슬롯마다 파악하여 프레임 크기를 동적으로 조정하는 Q-algorithm이라는 DFSA (Dynamic

* 이 논문은 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2018R1D1A1B07049849)

• First Author : (ORCID:0000-0003-3025-1161)Dept. Computer Engineering, Hanbat National University, dkrtpsxm123@gmail.com, 학생(석사), 정회원

^o Corresponding Author : (ORCID:0000-0001-6795-8086)Dept. Computer Engineering, Hanbat National University, gabriel@hanbat.ac.kr, 정교수, 종신회원

논문번호 : 202006-123-B-LU, Received June 9, 2020; Revised July 6, 2020; Accepted July 6, 2020

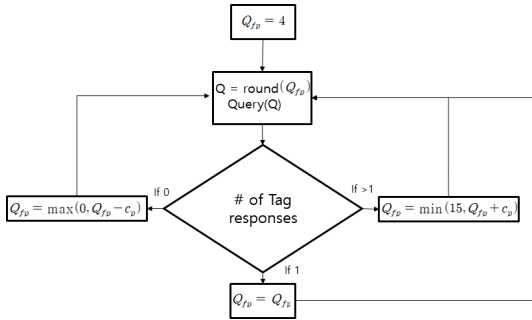


그림 1. Q-algorithm의 프레임 크기 결정
Fig. 1. Frame Size Adjustment by Q-algorithm

Frame Slotted Aloha)의 변형을 사용한다.

Q-Algorithm은 그림 1과 같이 동작한다. Q_{fb} 는 프레임 크기를 결정하는 매개변수로 EPC-C1G2에서는 2의 Q_{fb} 제곱을 프레임 크기로 사용한다. 각 라운드에 매 슬롯마다 전송 태그의 수가 1개일 경우 Q_{fb} 의 크기는 변하지 않고, 0개일 경우 c만큼 감소하고, 2개 이상일 경우는 c만큼 증가한다. 이때 c의 범위는 0.1보다 크고 0.5보다 작은 범위로 설정할 수 있다. 이 과정을 진행하면서 프레임을 언제 종료하는지에 따라 2가지 방식으로 나뉜다. 프레임의 끝까지 진행한 후에 Q_{fb} 의 반올림 값을 사용하는 것이 EPC-Q-Frame이고, Q_{fb} 이 직전의 프레임의 Q_{fb} 와 다를 때 프레임을 끝까지 진행하지 않고 조기에 프레임을 마감하고 다시 프레임의 크기를 설정하여 다음 프레임으로 넘어가는 것을 EPC-Q-Slot이라고 한다.

2.3 RL-DFSA (Reinforcement Learning DFSA)

RL-DFSA는 DFSA 방식인 Q-algorithm을 강화학습의 한 종류인 Q-Learning을 통해 성능을 개선하였다⁴⁾. 상태(State)는 한 프레임에서 충돌이 발생한 슬롯의 개수로 정의하고 행동(Action)은 11개의 프레임 크기 변화로 정의하였다. 또한 보상(Reward)은 프레임의 전체 슬롯 중 충돌이 발생한 슬롯의 비율을 4단계로 정의하여 사용하였다.

RL-DFSA에서는 행동으로 사용되는 프레임 크기 결정 방식으로 남아있는 태그 개수에 1.46배를 하면 가장 적절한 프레임 크기로 사용할 수 있다는 [2]의 결과를 사용하여 11개의 행동으로 정의했다. RL-DFSA에서는 학습을 할 때 태그 수를 1000개로 설정하여 충돌이 발생할 수 있는 경우인 1~500까지 총 500개의 상태가 존재한다.

III. 제안하는 강화학습 기반 충돌회피 알고리즘

본 논문에서는 RL-DFSA를 변형한 형태로 성능을 높이는 방안을 제안한다. RL-DFSA는 상태를 충돌이 발생한 슬롯의 개수로 설정하기 때문에 학습이 적은 상태가 발생할 가능성이 높아 Throughput이나 태그 전송 완료 시간에 영향을 준다. 제안된 방식은 상태의 개수를 줄이고 성능을 높이기 위해 프레임의 슬롯 개수와 충돌이 발생한 슬롯의 비율을 상태로 사용한다. 상태는 비율을 1% 구간으로 나눠서 101 (0%~100%)개로 구성하였다. 행동은 11개로 이전 프레임의 충돌 발생 슬롯의 개수(C)를 기반으로 식 (1)과 같이 사용하였다.

$$Action(i) = C \times 1.46 \times (2.0 + 0.2 \times i) \quad (1)$$

보상값(reward)은 현재 프레임 크기의 성공한 슬롯의 비율에 따라 다른 값을 주어 학습한다. 제안한 방식에서 사용한 보상은 충돌이 발생한 범위에 따라 달라지는데 그 값은 표 1과 같다.

그림 2는 Q-learning 기법을 사용한 DFSA 기법을 나타낸다. 학습을 할 때 C(충돌이 발생한 슬롯의 수), S(1개의 태그만 들어온 슬롯의 수), E(태그가 들어오지 않은 슬롯의 수)를 초기화 한다. 프레임 크기를 태

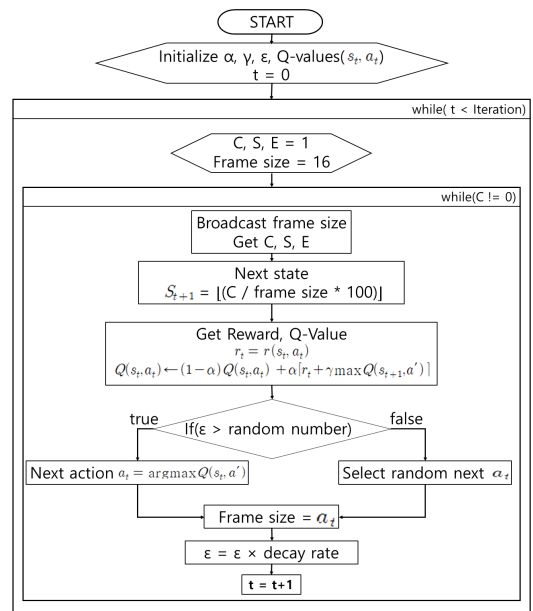


그림 2. Q-learning 기반의 충돌회피 알고리즘
Fig. 2. Anti-collision algorithm based on Q-learning

표 1. 제안된 방식에서 사용한 보상값
Table 1. Reward used in the proposed scheme

Rate of Success Slots	Reward
$0 \leq \text{rate} < 0.2$	1
$0.2 \leq \text{rate} < 0.4$	2
$0.4 \leq \text{rate} < 0.6$	4
$0.6 \leq \text{rate} < 0.8$	8
$0.8 \leq \text{rate} < 1.0$	16

그들에게 알리고 태그들의 전송을 수신한 후 C, S, E 를 얻고 다음 상태인 S_{t+1} 를 확정한다. 상태에 따른 보상을 그림 2의 값을 사용하여 결정하고 Q값을 업데이트한다. 0부터 1사이의 임의의 값을 설정하고 이 값이 ϵ 보다 크면 가장 큰 Q값을 가지는 행동을 통해 다음번 프레임 크기를 결정하고 임의의 값이 ϵ 보다 작으면 행동 중에서 임의로 선택한다. 이 과정을 C가 0이 될 때까지 반복하면 1번의 학습이 종료된다. 학습이 종료될 때마다 ϵ 은 일정한 비율만큼 감소하며, 이런 과정이 반복되면서 무작위로 선택되는 행동의 비율이 감소한다. 성능평가에서는 표 2의 파라미터를 적용했다.

표 2. 성능평가에서 사용한 파라미터
Table 2. Parameters used in performance evaluations

Parameters	Value
Initial state	10
Number of actions	11
Learning rate, α	0.1
Discount rate, γ	0.9
Exploration, ϵ	0.3
Epsilon decay rate	0.99971
Iteration	30,000
Number of tags	1,000
Initial frame size	16
Duration of successful slot T_s	1.1ms
Duration of empty slot T_e	113.97 μ s
Duration of collision slot T_c	223.11 μ s

IV. 성능 비교

시뮬레이션을 통해서 제안한 방식과 기존 방식들의 성능을 비교하였다. 성능 평가 기준은 처리율과 태그 전송 완료 시간을 사용하였다. 처리율은 수식 (2)와 같이 계산하였다. T_s 는 성공 슬롯의 시간, T_e 은 빈

슬롯의 시간, T_c 은 충돌 슬롯의 시간, N_s 은 전체 성공 슬롯의 개수, N_e 은 전체 충돌 슬롯의 개수, N_c 은 전체 빈 슬롯의 개수를 나타낸다.

$$\text{Throughput} = \frac{N_s \times T_s}{N_s \times T_s + N_e \times T_e + N_c \times T_c} \quad (2)$$

그림 3은 제안된 방식과 이상적인 경우, EPC 표준, RL-DFSA 방식에 대한 처리율의 값을 비교한 그래프이다. Ideal 경우는 리더가 남아있는 태그의 개수를 정확하게 알고 있다고 가정하는 경우로 남은 태그의 수에 1.46배를 사용한 것이고 EPC 표준은 c값을 0.3을 사용할 때의 EPC-Q-Slot 알고리즘을 사용하였다. 처리율은 Ideal의 경우 대부분 80% 정도를 유지하고, 표준인 Standard(0.3)의 경우 76%~77%의 결과를 보여주고 있다. RL-DFSA의 경우 표준과 비슷한 결과를 나타내고, 제안된 방식은 기존의 표준이나 RL-DFSA보다 더 높은 처리율을 보여준다.

그림 4는 태그의 개수에 따른 전송 완료 시간이다. 처리 시간을 구하는 방법은 수식(3)과 같다.

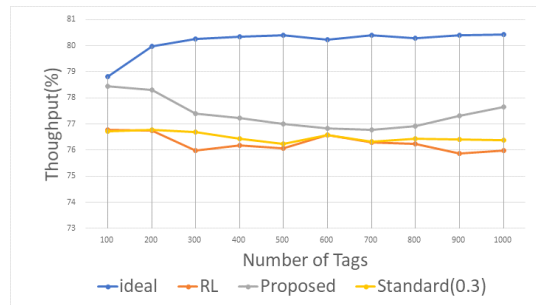


그림 3. 태그 개수에 따른 처리율
Fig. 3. Throughput vs number of tags

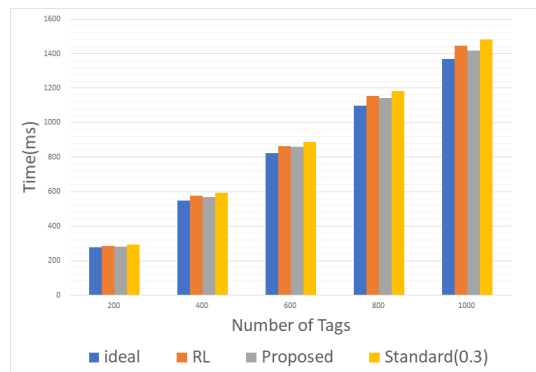


그림 4. 태그 개수에 따른 전송 완료 시간
Fig. 4. Transmission completion time vs number of tags

$$N_s \times T_s + N_e \times T_e + N_c \times T_c + N_R \times T_{query} \quad (3)$$

태그 전송 완료 시간은 각 슬롯의 종류별 시간과 개수를 곱한 값과 총 진행된 라운드의 개수(N_R)에 Query의 시간(T_{query})을 곱한 값을 더한 것이다. 제안된 방식이 Ideal을 제외한 RL-DFSA와 EPC 표준 방식보다 전송 완료 시간이 적게 걸리는 것을 확인할 수 있다.

V. 결 론

본 논문에서는 강화학습을 기반으로 한 RFID의 충돌 회피 알고리즘을 제안하여 기존 방식대비 성능이 향상된 것을 보여주었다. 강화학습의 특성상 초기상태와 행동, 보상값들이 성능에 영향을 주기 때문에 추후 파라미터 최적화 방안에 관한 연구가 필요하다.

References

- [1] J. R. Cha and J. Kim, "ALOHA-type anti-collision algorithms using tag estimation method in RFID system," *J. KICS*, vol. 30, no. 9, pp. 814-821, Sep. 2005.
- [2] S. Dhakal and S. Shin, "Precise-optimal frame length based collision reduction schemes for frame slotted aloha RFID systems," *KSII Trans. Internet Inf. Syst.*, vol. 8, no. 1, pp. 165-182, Jan. 2014.
- [3] EPCglobal, *Specification for RFID Air Interface EPC Radio-Frequency Identity Protocols Class-1 Generation-2 UHF RFID*, 2008.
- [4] M. Loganathan, T. Sabapathy, M. E. Elshaikh, and M. N Osman, "Reinforcement learning based anti-collision algorithm for RFID systems," *Int. J. Comput.(CIS)*, pp. 155-168, Jun. 2019.