[[1]](#footnote-1)

차량 통신 네트워크에서 다중 에이전트 Deep Q-Network를 이용한 자원 할당 기법

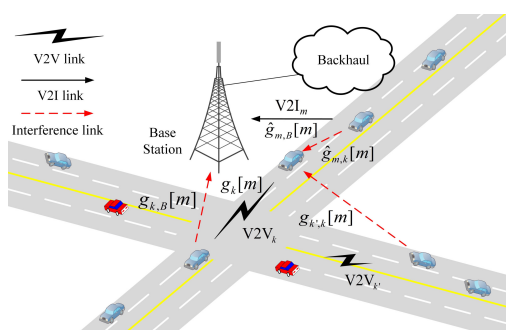
Donghee Han

***요약*—본 논문은 차량 통신 네트워크에서 다중 에이전트 Deep Q-Networks(DQN)을 적용한 분산형 자원할당 기법을 제안한다. 차량들은 공유 자원 리소스 블록(RB)과 전송 전력 신호를 선택하여 신호를 전송한다. 이 과정에서 차량 간의 간섭 신호를 제어하면서 Signal-to-Interference-plus-Noise Ratio(SINR)를 증가시키는 것은 중요하다. 제안하는 다중 에이전트DQN 기반 분산형 자원할당 기법은 Vehicle to Vehicle(V2V) 링크, Vehicle to Infrastructure(V2I) 링크의 채널 상태 정보와 전송 제약 시간을 고려하여 최적의 RB와 전송 신호 전력을 선택하는 것을 목적으로 한다. 모의실험은 V2I, V2V 링크의 데이터 처리율과 V2V 링크의 통신 실패 확률을 평가한다.**

# I. 서론

V2X 통신은 근접한 차량 간의 협동 시스템을 지원하여 교통 효율과 도로 안전성을 향상시키는 핵심기술이다[1]. 차량 통신 환경에서의 자원할당 시스템은 다양한 방법으로 제시되어 왔다. 본 논문에서 제안하는 DQN을 적용한 분산형 자원할당기법을 통해 V2V, V2I 링크의 상호 간섭을 최소화하고 V2V 링크의 전송 제약 시간을 충족하는 알고리즘을 제안한다.

# II. 시스템 모델

그림 1과 같이 차량 통신 시나리오에서 차량들이 V2V, V2I 링크로 통신하는 상황을 고려한다. V2I 링크를 사용하는 대의 차량 집합은 로 표시하고, V2V 링크를 사용하는 대의 차량 집합은 로 표시한다. V2I 링크를 사용하는 번째 차량의 SINR은  **그림. 1.** 차량 통신 네트워크

다음과 같다.

(1)

그리고 V2V 링크를 사용하는 번째 차량의 SINR은 다음과 같다.

(2)

이때 은 번째 V2I 링크의 전송 전력, 번째 V2V 링크의 전송 전력이고, 은 노이즈 신호이다. ]은 기지국과 형성된 RB의 채널 상태이다. 은 번째 V2V 링크가 동일한 RB를 선택한 경우 1, 아닌 경우 0인 이진 값을 가진다. 은 번째 V2V 링크의 채널상태 정보이다. 는 번째 차량의 전송 신호 전력이다. 은 번째 V2V 링크의 채널상태 정보이다. 그리고 의 공식은 다음과 같다.

(3)

번째 V2I 링크의 채널 용량 공식은 다음과 같다.

(4)

그리고 번째 V2V 링크의 채널 용량 공식은 다음과 같다.

(5)

여기서 는 대역폭이다. 위에서 언급한 공식에 따라 V2I 링크의 총 채널 용량은 다음과 같은 식을 따르게 된다.

(6)

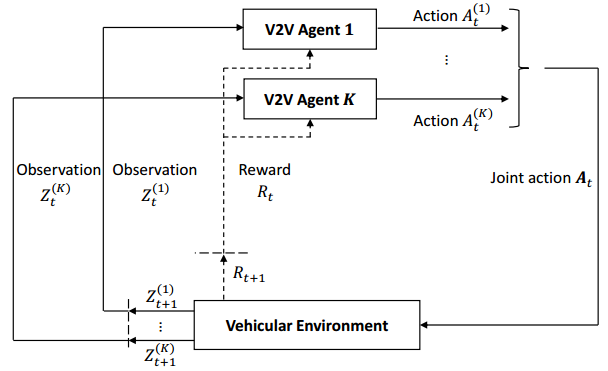
그리고 V2V 링크의 전송 제약 시간을 충족 확률식은 다음과 같다.

(7)

여기서 는 V2V 링크에서 요구하는 데이터 량이고, 는 V2V 링크의 제한 시간, 는 channel coherence time, 는 시간 슬롯에서 겪는 번째 V2V 링크의 채널 용량이다.

# III. 다중 에이전트 Deep-Q-Network기반 자원 할당

그림 1과 같은 차량 통신 환경에서 모든 V2V 링크는 고유의 에이전트를 할당 받는다. 에이전트는 현재 V2V 링크가 경험하는 튜플 데이터를 기반으로 학습한다. 학습 과정은 그림 2와 같다.



**그림. 2.** 다중 에이전트 DQN 학습 방법.

첫번째로 모든 V2V 링크는 환경으로부터 지역적인 데이터를 관측한다. 관측 후 V2V 링크들은 에이전트를 기반으로 행동을 결정한다. 결정한 행동들은 모든 V2V 링크가 결정하기까지 수집된다. 모든 V2V 링크가 결정을 완료하면 환경으로 모든 행동을 가한다. 그 다음 환경으로부터 보상을 얻는다. 보상은 1개의 스칼라 값이다. 따라서 모든 에이전트들은 각기 다른 행동을 했지만 모두 같은 보상을 받게 된다. 이와 같은 절차를 V2V 링크의 에이전트들은 학습하여 결과적으로 상호 협력적인 자원할당 기법을 하는 것을 목표로 한다.

## A. 상태 공간

다중 에이전트 DQN의 상태 공간은 다음과 같이 정의했다.

여기서 는 번째 V2V 링크의 남은 데이터 량,

은 번째 V2V 링크의 남은 시간, 은 모든 RB에서 겪는 간섭 신호, 은 번째 V2V 링크의 채널 상태, 는 다른 V2V 링크에 의한 간섭 신호, 는 기지국에 의해 겪는 간섭 신호, 는 모든 V2I 링크의 채널 상태 정보이다.

## B. 행동 공간

행동 공간은 개의 선택 가능한 RB와 전송 전력 레벨을 [23, 10, 5, -100]dBm을 제어할 수 있다. 따라서 행동 공간 은 크기를 가진다.

## C. 보상

다중 에이전트 DQN의 모든 에이전트들은 같은 보상을 받기 때문에 전역적인 영향을 받은 보상이어야 한다. 따라서 모든 V2I 링크의 채널 용량 합과 모든 V2V 링크의 채널 용량 합 또는 상수를 지정했다. 보상함수는 다음과 같다.

, (9)

, (10)

여기서 는 보상 함수 가중치이다. 는 V2V 링크가 제한 시간안에 요구한 데이터 량을 충족했을 때 얻는 보상 상수로 지정한다. 이때 값은 V2V 링크들의 채널 용량 합 값보다 높아야 한다.

## D. 학습

다중 에이전트 DQN을 학습시키기 위해서 각 에이전트들은 독립적으로 메모리를 구성하였다. 독립적으로 구성한 메모리는 다음과 같은 튜플 데이터로 저장된다.

, (11)

여기서 는 t 시간 슬롯에서 에이전트가 선택한 행동을 의미한다. 그리고 에피소드가 끝날 때 배치 사이즈만큼 메모리부터 튜플 데이터를 추출하여 에이전트의 신경망을 학습시킨다. 학습은 다음 공식을 최소화하는 것을 목표로 학습한다.

, (12)

여기서 는 대상 Q-network, 는 현재 학습 중인 Q-network를 의미한다.

표 I

모의 실험 파라미터

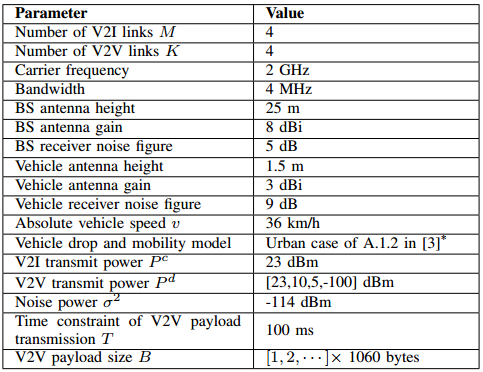
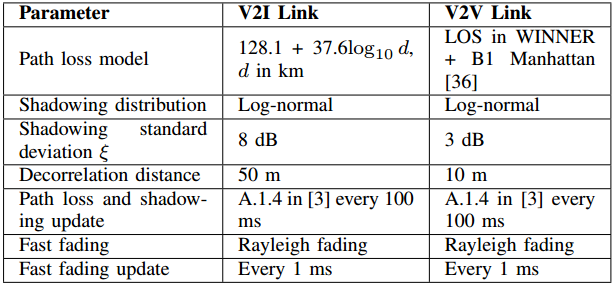


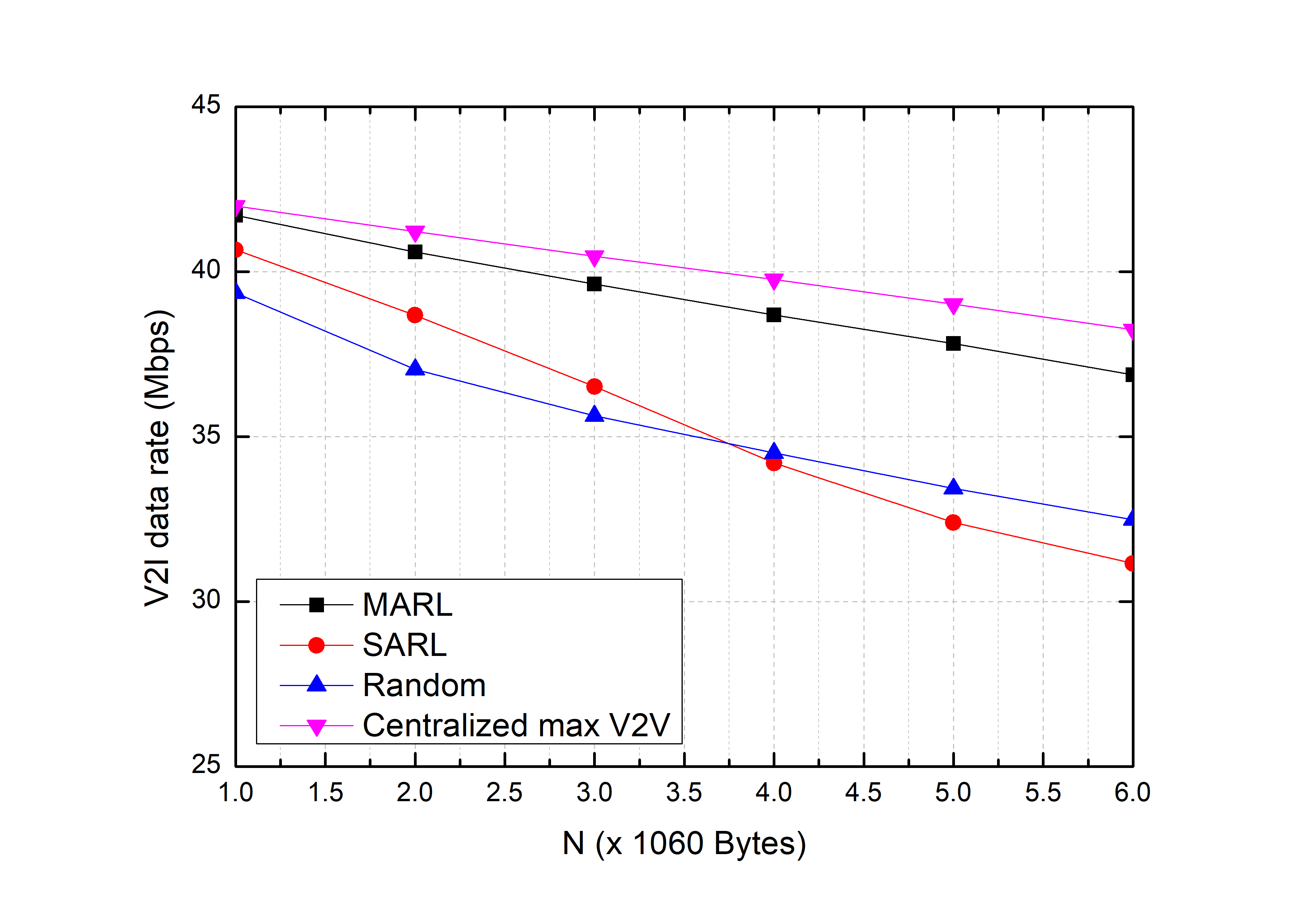
표 II

모의 실험 채널 모델

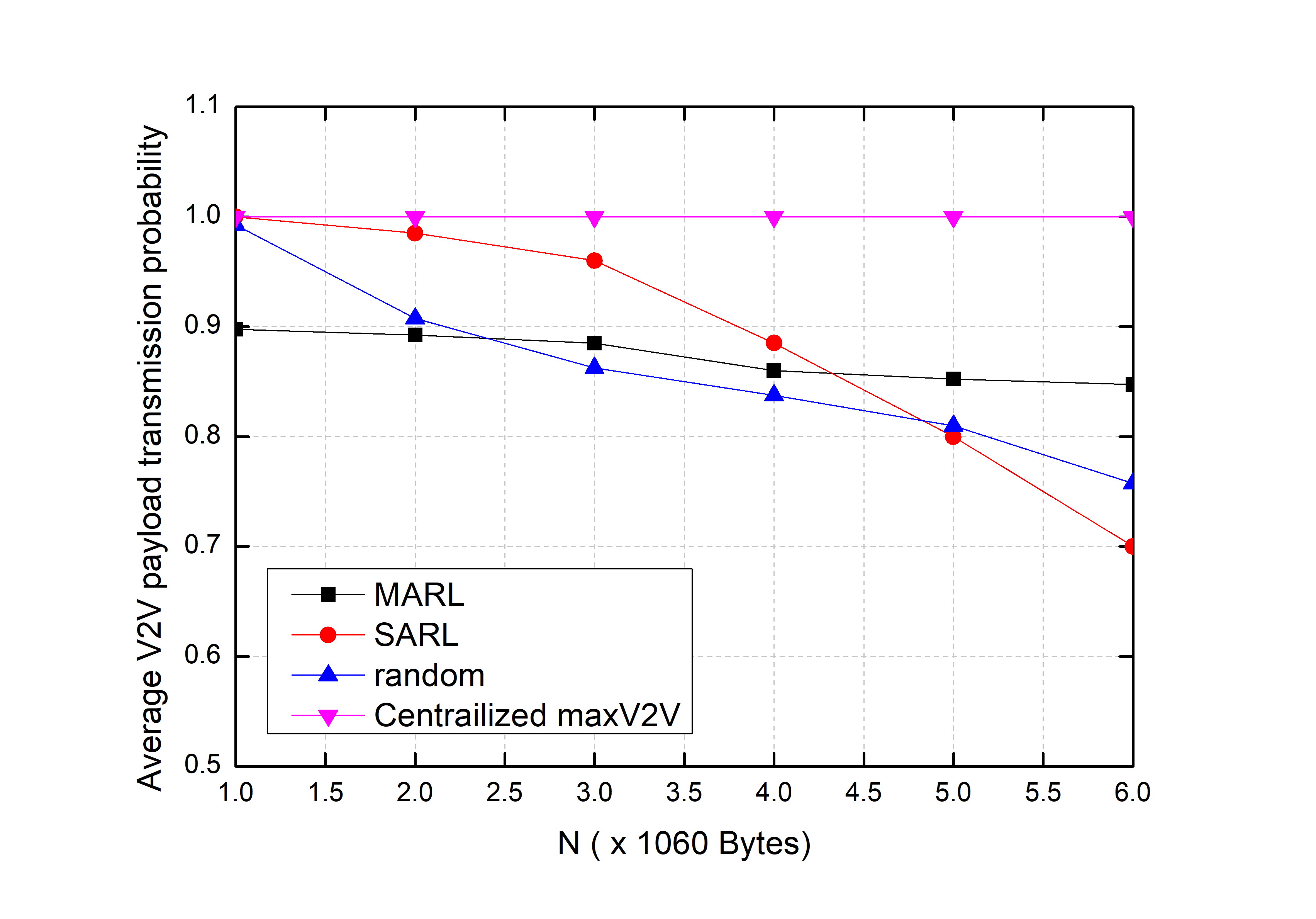


# IV. 모의 실험

모의 실험은 3GPP TR 36.885의 Urban case을 구성하여 실험하였다.



**그림. 3.** Average sum rate of V2I links

 **그림. 4.** Average V2V payload transmission probability

실험그래프는 제안하는 다중 에이전트 DQN, 단일 에이전트 DQN, 무작위로 선택하는 “Random” 방법, 모든 선택 방법을 고려한 방법으로 비교하였다. 요구하는 V2V 링크의 데이터 량이 늘어날수록 단일 에이전트 DQN, Random 방법보다 더욱 좋은 성능을 보이고 있다. 하지만 요구하는 데이터 량이 적으면 다중 에이전트 DQN이 상대적으로 성능이 떨어지는 것을 볼 수 있다.

# V. 결론

## 본 논문은 V2X 차량 통신 시나리오에서 다중 에이전트 DQN을 적용한 자원 할당 기법을 제안하고 성능을 평가하였다. 제안하는 다중 에이전트 DQN 기반 전송 전력과 단일 에이전트 DQN, 무작위 기법과 비교하였을 때 처음에는 성능이 떨어지나 V2V 링크의 데이터 량이 늘어날수록 다른 기법 대비 높은 성능을 보였다. 그러나 다중 에이전트 DQN 기법은 행동 공간이 이산적인 문제가 있으므로 추 후 DDPG와 같은 강화학습 기법을 이용해 연속적인 행동 공간에서 자원 할당 기법을 연구할 계획이다.

참고 문헌

# L. Liang, H. Ye, and G. Y. Li, ““Spectrum sharing in vehicular networks based on multi-agent reinforcement learning,” *IEEE J. Sel. Areas Com-* *mun., vol. 37, no. 10, pp. 2282–2292, Oct. 2019.*

1. H. Ye, G. Y. Li, and B. F. Juang, “Deep reinforcement learning based resource allocation for V2V communications,” IEEE Trans. Veh. Technol., vol. 68, no. 4, pp. 3163–3173, Apr. 2019.
2. T. S. G. R. A. Network, “Study LTE-based V2X services; (release 14),” document 3GPP TR 36.885, vol. V14.0.0, June 2016.

1. [↑](#footnote-ref-1)