2003년도 대한전자공학회 하계종합학술대회 제26권 제1호

Q-Learning 을 이용한 Intelligent Transportation System

박명수, 김표재, 최진영

서울대학교 전기컴퓨터공학부 지능연산제어연구실

전화: 02-872-7283

Intelligent Transportation System using Q-Learning

Myoung Soo Park, Pyo Jae Kim and Jin Young Choi

School of Electrical Engineering & Computer Science,

Seoul National University, Seoul, Korea

E-mail:mspark@neuro.snu.ac.kr, pjkim@neuro.snu.ac.kr, jychoi@neuro.snu.ac.kr

Abstract

In this paper, we propose new method which can provide user the path to the target place efficiently. It stores the state of roads to target place as the form of Q-table and finds the proper path using Q-table. Q-table is updated by the information about real traffic which is reported by users. This method can provides the proper path, using less storage and less computation time than the conventional method which stores entire road traffic information and finds the path by graph search algorithm.

I. 서론

현대 사회가 복잡해 짐에 따라 교통 체증으로 인한 자원 및 시간의 낭비, 오염 등이 사회적 문제가 되고 있다. 이러한 교통 체증으로 인하여, 목적지까지 원하는 시간에 도착하도록 출발지에서 목적지까지의 최적의 경로를 찾는 것이 중요한 문제로 부각되었다. 이 문제를 해결하기 위한 기존의 방법은 모든 도로의 교통 상황을 저장하고 graph search 알고리즘을 이용하여 최적의 경로를 찾아내는 것이다. 그러나 이러한 graph search 방법은 교통 체계가 복잡해 짐에 따라 적용하기 어려운 단점을 가지고 있다. 일례로 graph search 알고리즘은 소수의 사용자만이 이용하거나 아니면 이용하는 사용자가 없는 도로의 교통 정보까지 알고 있어야 최적의 경로를 계산할 수 있으며 출발지에서 목적지까지의 가능한 모든 경로를 계산한 후에야 최적의 경로를 결정할 수 있게 된다. 이로 인하여 복잡한 교통 체계에서는 최적의 경로를 계산하기 위해 좀더 많은 계산과 저장 공간을 필요로 하게 된다.

이에 대한 해결 방법으로서 본 논문에서는 Q-learning을 이용한 ITS (Intelligent Transport System)을 제안하고자 한다. Q-learning을 이용하여 좀더 적은 저장 공간과 계산량으로 사용자에게 목적지까지의 경로를 제공하는 새로운 방법을 제시하고 이를 검증하기 위하여 가상의 교통망을 이용한 실험 방법을 제시하도록 하겠다.

II. Q-learning 을 이용한 Intelligent Transports System (ITS)

기존의 ITS에 많이 사용되어 왔던 graph search 알고리즘은 최적의 경로를 찾기 위하여 모든 도로의 교통 상황에 대한 정확한 정보를 알고 있어야 한다. 따라서 교통 체계의 규모가 커지고 복잡해 짐에 따라 이 들 정보의 저장을 위한 공간이 많이 필요하게 되며, 모든 경우에 대한 경로를 고려해보아야 하기 때문에 계산 시간이 많이 걸리게 된다. 이는 경로 정보 제공자에게 많은 부담을 주게 된다.

보통 ITS의 주된 목적지는 많은 사람 들이 이용하는 백화점이나 극장, 공항, 버스 터미널, 테마파크와 같은 대형 장소 들이며 이 들에 대하여 별개의 ITS를 구성하여 서비스 한다면 경로 정보 제공자의 부하를 분산 시킬 수 있다. 즉 사용자는 목적지까지의 경로를 목적지에 설치된 ITS에게 요구함으로써 얻게 되는 것이다. 또한 최적의 경로를 찾기 위해 기존의 graph search 알고리즘 대신에 Q-learning을 사용함으로써 경로 계산을 위한 시간과 저장공간의 절약이 가능하다.

이 Q-learning을 이용하는 새로운 ITS의 기본 구조는 그림 1과 같다.

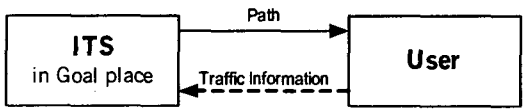

그림 1. ITS using Q-learning

사용자가 임의의 목적지까지의 경로를 요청하면, 목적지에 설치된 ITS는, Q-table에서 현재 위치에서 이동 가능한 방향 중에 가장 작은 Q-value를 가지는 경로를 제공하게 된다. 이 때 경로 선택의 기준이 되는 Q-value는 목적지까지의 소요 시간을 반영하므로 작은 Q-value를 가지는 경로를 따라가면 빠른 시간에 목적지까지 도착하게 된다. 사용자는 제공된 경로가 이용할 수 없거나 다른 경로가 더 좋다고 생각되면 다른 경로를 이용할 수 있으며 그렇지 않다면 ITS로부터 제공된 경로를 따라가게 된다. 사용자는 목적지에 도착한 이후에 이용한 경로에 대한 교통 정보를 ITS에 전달하며 ITS는 전달된 교통 정보를 바탕으로 Q-learning을 이용하여 Q-value가 이를 반영하도록 업데이트하게 된다. 만약 지금 도착한 사용자가 거쳐온 경로의 도착 시간이 Q-value를 이용해서 계산되었던 예상 도착 시간 보다 짧다면 Q-learning은 보상과 함께 다음 식에 따라 Q-value를 업데이트 하게 된다. [1]

$$\hat{Q}(s,a) \leftarrow r + \gamma \max_{a'} \hat{Q}(s',a') \quad s \leftarrow s'$$

여기서 $\hat{Q}(s,a)$ 는 현재 상태 s 에서 취할 수 있는 행동 a 에 대항 Q-value이며 r 은 행동에 대한 보상 값이다. 이러한 방법으로 ITS는 Q-learning을 이용하여 실제 교통 상황을 반영할 수 있도록 Q-table을 업데이트한다. 그림 2는 ITS에서 일어 나는 Q-learning의 간략한 블록도 이다.

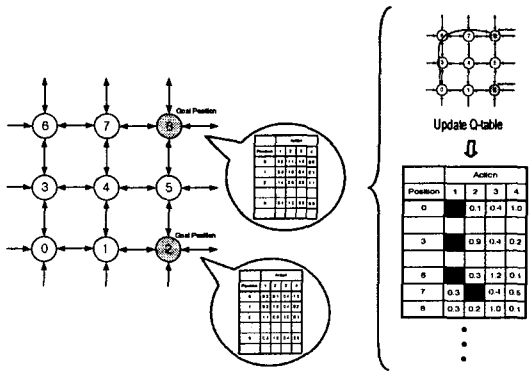

그림 2. Q-learning 블록도

Q-learning을 이용한 새로운 ITS는 다음과 같은 두 가지 장점을 가진다.

첫째로, 사용자들에 의해 거의 사용되지 않은 도로에 대한 정보를 저장하거나 업데이트할 필요가 없다. ITS는 목적지에 도착하기 위해 사용자들에 의해 사용되는 경로에 대한 교통 정보만을 저장하고 업데이트하게 된다. 이로 인하여 큰 저장 공간을 필요로 하지 않으며 현재 교통 상황을 반영하기 위한 업데이트에 걸리는 시간이 짧다.

둘째로, Q-learning을 이용함으로써 기존의 graph search 알고리즘을 이용하는 방식에 비하여 보다 적은 계산량을 필요로 한다. 기존의 방법은 현재 위치에서 목적지까지의 가능한 모든 경로에 대한 탐색이 필요하지만 새로운 ITS에서는 작은 Q-value를 가지는 경로만은 찾으면 되는 것이다.

Q-learning을 사용한 ITS에 의해 제공된 경로는 사용자들에 의해 잘 사용되지 않는 도로에 대한 정보는 반영되지 않으므로 목적지까지의 최단 경로가 되는 것은 아니다. 그러나 사용자들에 의해 이러한 도로에 대한 정보가 제공되게 되면 Q-learning에 의해 제공되는 경로는 최적의 경로에 가까워지게 되며, 보다 적은 계산으로 비교적 정확하며 빠른 경로를 계산할 수 있게 된다.

III. 가상 교통망을 이용한 실험

Q-learning 을 이용한 ITS 의 효용성을 검증하기 위하여 다음과 같은 가상 교통망을 구현하였다.

가상 교통망에는 미리 설정된 임의의 개수의 목적지가 존재하며 각 목적지에는 Q-table 을 가지는 ITS 를 구현하였다. ITS 서비스를 이용하는 사용자들은 임의로 설정된 자신의 목적지까지의 빠른 경로를 ITS 에 요청하고 ITS 는 현재까지 학습된 Q-table 에 근거하여 목적지까지의 경로를 제공하게 된다. 또한 ITS 를 사용하지 않은 비사용자들이 존재함으로써 실제 상황 교통 상황과 비슷한 상황이 발생하도록 하였다.

이 교통망을 이용하여, 새로운 ITS 의 효용성을 검증하기 위한 두 가지 실험을 할 수 있다.

첫째로, 동일한 목적지를 가지는 ITS 를 이용한 사용자와 비사용자들의 소요시간을 비교하도록 한다. 실험에 사용되는 ITS 비사용자들은 미리 설정된 경로(목적지까지 가는 가장 일반적인 경로)를 따라 동일하게 움직이게 된다. 그에 반해 ITS 를 이용하는 사용자들은 목적지에 도착 후 자신의 경로에 해당하는 교통 정보를 ITS 에 제공하며 ITS 는 이에 근거하여 Q-value 를 업데이트하게 된다. 우선 ITS 의 Q-table 이 가상 교통망의 교통 상황을 반영하도록 하기 위하여 일정 시간의 학습 시간을 갖게 된다. 초기의 Q-value 들은 임의의 큰 값으로 설정되며 일정 시간의 학습이 진행되면 Q-value 는 가상 교통망의 교통 상황을 반영하게 된다. 이때 ITS 에 의해 제공되는 경로의 소요 시간을 측정하여 이와 비교함으로써, 학습결과가 축적됨에 따라 교통상황이 반영되는 것을 확인할 수 있

둘째로 기존의 graph search 알고리즘을 이용한 ITS 에 대하여 Q-learning 을 이용하는 ITS 의 효율성을 보이기 위하여 동일한 가상 교통망 상황하에 두 방법을 적용한 실험을 수행하도록 한다. 최적의 경로를 제공하기 위하여 사용된 저장공간과 계산 시간 그리고 목적지까지의 소요 시간을 비교하여 Q-learning 을 사용한

ITS 의 효용성을 확인할 수 있다.

IV. 결론

본 논문에서는 Q-learning 을 이용한 새로운 Intelligent Transportation System 을 제안하였다. 이 새로운 방법은 기존의 방법에 비해 보다 적은 저장 공간과 계산을 통해 목적지까지의 경로를 제공할 수 있다. 두 가지 실험 방법을 통해 이를 검증할 수 있다.

References

[1] Tom M. Mitchell, "Machine Learning," 1997, McGraw Hill