

# 다중 교차로환경에서 강화학습기반 분산 교통신호 제어

주 현 진\*, 임 유 진<sup>o</sup>

## Distributed Traffic Signal Control at Multiple Intersections Based on Reinforcement Learning

Hyunjin Joo\*, Yujin Lim<sup>o</sup>

요 약

지능형 교통 시스템(ITS)은 스마트 시티에서 효율적으로 교통관리를 가능하게 해주는 시스템이다. 교통관리 시스템은 교통 체증을 해결하기 위해 다양한 방식으로 연구되어왔는데, 가장 일반적으로 사용되는 시스템은 교통신호 제어 시스템이다. 과거에는 각 교차로 별로 교통신호 제어를 고려하였다. 하지만 성능의 최적화를 위해서 최근에는 주변 교차로를 모두 고려한 교통신호 제어 기법을 많이 연구하고 있다. 본 논문에서는 Q-learning 강화학습을 이용하여 다중 교차로 환경에서 교통신호 제어 기법을 제안한다. 본 제안기법의 목적은 교차로에서의 처리량을 최대화하고, 인근 교차로의 대기시간을 최소화하는 것이다.

**키워드** : 교통신호 제어, 강화학습, 분산 방식, 처리량, 대기 시간

**Key Words** : traffic signal control, reinforcement learning, distributed control, throughput, waiting time

### ABSTRACT

Intelligent Transportation System(ITS) is a system for efficient traffic management in smart cities. The traffic management system has been studied in various ways to solve traffic congestion, the most commonly used system is a traffic signal control system. In the past, traffic control was considered for each intersection. However, in order to optimize performance, traffic signal control techniques have recently studied considering all surrounding intersections. In this paper, a system for controlling traffic signals at multiple intersections is proposed using Q-learning which is reinforcement learning. The purpose of this study is to maximize throughput at intersections and minimizing waiting time at adjacent intersections.

### I. 서 론

스마트시티는 도시의 인구 집중, 기반 시설 노화, 에너지 부족, 교통 혼잡 등 다양한 도시 문제를 해결하기 위한 주목받는 미래 동향 중 하나이다. 스마트시티는 사물인터넷(IoT), 빅데이터(big data), 인공지능(artificial intelligence), 자율주행차 등 다양한 혁신기

술을 도시 인프라와 결합하여 다양한 도시 서비스를 제공한다<sup>1)</sup>. 스마트시티는 표 1과 같이 스마트 하우스, 스마트 교통, 그린시티(green city), 스마트시티 관리, 스마트 관광 등으로 다양한 분야로 구성되어있다. 스마트시티와 관련된 전 세계의 프로젝트 중, 50% 이상이 교통 부문과 관련이 있다. 이는 교통이 스마트시티에서 다루는 요소들 중에서 중요한 부분을 차지하고

\* 이 성과는 2018년도 정보(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2018R1A2B6002505)

• First Author : Sookmyung Women's University Department of IT Engineering, hjjoo8@sookmyung.ac.kr, 학생회원

o Corresponding Author : Sookmyung Women's University Department of IT Engineering, yujin91@sookmyung.ac.kr, 종신회원  
논문번호 : 201911-321-B-RN, Received November 27, 2019; Revised December 4, 2019; Accepted December 4, 2019

있다는 것을 증명한다.

지능형 교통 시스템(intelligent transport systems)에는 교통신호 관리, 교통 최적화, 주차 관리 등 교통과 관련된 서비스를 제공하는 차세대 교통 시스템이다<sup>2)</sup>. 이는 스마트시티 안에서 효율적인 도로 교통관리를 가능하게 만들고 운전자들에게 안전과 편리성을 제공한다. 그 중, 교통관리 시스템(traffic management system)은 CCTV 또는 센서(sensor)로부터 정보를 수집하여 실시간 교통흐름을 관리하고 제어할 수 있다. 이는 교통관리 체계의 효율성을 높인다.

교통 체증은 많은 도시 지역에서 매일 발생하는 교통 문제이다. 교통 체증을 해결하기 위해 교통관리 시스템은 다양한 방식으로 연구되어왔다. 교통신호 제어(traffic signal control)는 가장 일반적으로 사용되는 교통관리 시스템이다. 현재 사용되고 있는 교통신호 시스템은 고정형 신호(fixed-time signals)를 이용한다. 고정형 신호는 사전에 입력한 몇 개의 신호시간에 따라 매일 반복하여 신호를 제어하는 방식이다. 하지만 고정형 신호 제어는 계속해서 변하는 교통 환경에 대해 효율적과 유동적으로 신호를 제공할 수 없다. 이전의 연구된 교통신호 제어 시스템은 수학적 모델(mathematical model)이나 최적화 기법(optimization technique)을 사용했다<sup>3,4)</sup>. 하지만, 수학적 모델과 최적화 기법을 사용한 교통신호 제어 시스템은 동적이고 대규모 데이터를 다루는데 있어서 복잡성이 높다. 이에 따라, 최근에는 교통신호 제어 시스템에 인공지능 기술을 적용하고 있다. 본 연구는 인공지능 기술 중 Q-learning(QL) 강화학습을 적용한 교통신호 제어 시스템을 제안하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 신호 제어 시스템과 관련된 연구에 대해 살펴본다. 3장에서는 제안하는 QL 기반 알고리즘에 대해 알아본다. 4장에서는 제안하는 알고리즘의 실험결과와 그에 대한 분석을 기술하고, 5장에서는 결론 및 향후 계획을 제시한다.

## II. 관련 연구

현재, 교차로 교통관리 연구는 다양한 방법으로 진행되고 있으며 그 중에서도, 교통신호 시스템에 관한 연구가 활발히 진행되고 있다. 현재 사용되는 신호 시스템은 고정형 방식이다. 고정형 방식보다 교차로의 처리량 늘리기 위해 적응형(adaptive) 방식도 연구되었다. 적응형 방식은 교통량에 따라 교통신호의 시간을 조절하거나 신호의 순서를 바꾼다.

표 1. 스마트시티 산업분야 및 관련 기술  
Table 1. Smart city industry and related technologies.

Industry	strategic items and technologies
Smart transportation	traffic management system, public transport system, electronic payment system, intelligent vehicle/road, intelligent parking assist system, etc.
Smart safety	real time disaster management system, intelligent CCTV control system, safety robot, safety center, safety alarm system, etc.
Smart environment	hazardous environment monitoring device, etc.
Smart energy	building energy management system, renewable energy integration control/management, etc.
Smart healthcare	mobile diagnostic device, real time healthcare system, medical information management platform, telemedicine service, etc.
Smart welfare	silver care monitoring service, muscle assisted robot for the ederly and the disabled, chronic disease monitoring service, etc.
Data intelligence	bigdata SW, cloud SW, AI SW, data security, image processing system, etc.
Information security	IoT security, drone security, healthcare/medical security, etc.

동적으로 변화하는 교통 환경에서 많은 양의 데이터를 사용하는 교통신호 최적화 문제는 기존의 수학적 모델 또는 최적화 방법으로 해결하기에 높은 복잡성의 문제가 발생한다. 교통신호 문제를 해결하기 위하여 퍼지(fuzzy) 기법과 QL기법 등을 많이 이용하고 있다.

단일 교차로에서 퍼지 기법을 이용하여 교통신호 제어 기술을 제안하였다<sup>5-6)</sup>. 퍼지 기법을 사용한 신호 제어는 녹색 신호의 순서는 고정되어 있으며 교통량에 따라 녹색 신호의 시간을 동적으로 조절한다. 교통량을 파악하기 위해 교차로에 유입되는 차량의 개수를 측정한다. Yang et al<sup>5)</sup>는 측정된 차량의 개수 정보를 이용하여 녹색 신호에서의 현재 교통흐름(traffic flow)과 다음 구간(next phase)에서 적색 신호 일 때의 교통흐름을 파악한다. 파악된 교통흐름에 따라 녹색 신호의 시간 연장 여부에 대해 결정한다. 즉, 녹색

신호 시간의 축소는 고려하지 않는다. 반면, Askerzada et al<sup>[6]</sup>은 측정된 차량의 개수 정보로 교통 흐름의 패턴을 결정한다. 결정된 패턴으로 녹색 신호의 시간을 줄이거나 확장한다. 퍼지 기법을 사용한 신호 제어는 동적인 교통 환경에서 고정형 신호보다 유연한 제어가 가능하다. 하지만, 퍼지 제어 모델은 환경이 변화함에 따라 퍼지 제어 룰(fuzzy control rule)이 변화하고 생성되기 때문에 오버헤드(overhead)가 크다. 따라서 강화학습인 QL을 사용한 교통신호 기술이 연구되고 있다<sup>[7-10]</sup>.

QL 기법은 강화학습 기법으로 학습(learning)을 통해 최적의 정책(policy)을 찾는다. QL의 장점은 미리 정의된 환경 모델을 사용하지 않는 것으로, 동적인 교통 환경에 적합하다. QL 기법을 이용한 교차로 신호 제어 연구는 크게 단일 교차로 단위 연구부터 다중 교차로를 함께 고려한 연구로 나눌 수 있다.

단일 교차로 연구는 단일 환경에서의 학습 경험을 얻고 다양한 매개변수의 유용한 범위를 설정하기 위해 진행된다<sup>[7-8]</sup>. 녹색 신호의 순서는 고정되어 있으며 학습에 따라 녹색 신호의 시간을 조절한다. Chin et al<sup>[7]</sup>은 매개변수로 큐 길이(queue length)를 고려하여, 항상 교차로에 차량이 가장 적게 있는 것을 목표로 가진다. 큐 길이는 교차로 레인(lane)에 차량이 대기하고 있는 길이이다. Abdulhai et al<sup>[8]</sup>은 매개변수로 큐 길이와 교차로를 빠져나간 시간을 고려하였다. 그리고 교차로 큐 길이의 편차를 줄여서 교통 지연(traffic delay)을 최소화하는 것을 목표로 한다. 단일 교차로에서의 신호 제어 연구도 중요하지만, 교차로는 연속되어있고 인접한(adjacent) 교차로에 영향을 받는 구조이다. 따라서 다중 교차로를 고려한 신호 제어 연구들이 진행되고 있다<sup>[9,10]</sup>.

다중 교차로 연구는 여러 교차로들의 상황을 고려해야하기 때문에 고려해야 할 변수가 다양하다. Khamis et al<sup>[9]</sup>은 다중 교차로 사이에서 여행시간(trip time)을 최소화하는 시스템을 제안한다. 여행시간이란 출발지부터 목적지까지 걸린 시간을 의미한다. 또한, 그린 웨이브(green wave)를 추구하는 시스템이다. 그린 웨이브란 진행하고 있는 차량 그룹을 대상으로 세 개 이상의 연속된 신호등을 점진적으로 켜서 교차로에서 멈출 필요 없이 지나가게 하는 상황을 뜻한다. 이는 차량 속도를 유지할 수 있기 때문에 연료 소비량을 최소화 할 수 있다. Chin et al<sup>[10]</sup>은 다중 교차로에서 녹색 신호의 시간을 연장하고 축소하는 방법으로 상황에 맞게 적절한 녹색 신호를 분배한다. 적응형 교통신호 기법으로, 큐 길이를 최소화하는 것을 목

표로 한다.

분석한 여러 교차로에 대한 비교논문들은 고정된 신호 순서에 따라 녹색 신호의 시간을 조절한다. 하지만, 교차로들은 연속적으로 구성되어있기 때문에 인접한 교차로들끼리 서로 영향을 받는다. 상황에 따라 신호의 순서가 유연하게 바뀐다면, 더 효율적으로 교통 신호를 제어할 수 있다.

본 논문은 분산방식을 사용한 교통신호 제어 시스템을 제안한다. 다중 교차로에서의 분산된 교통신호 제어를 위해서 본 연구는 각 교차로의 처리량과 인접한 교차로에서의 대기시간을 고려하여 QL기법을 이용한 새로운 교통신호 제어 시스템을 제안한다. 처리량은 단위 시간 동안 교차로를 지나간 차량의 개수이고, 대기 시간은 차량이 멈춘 시점부터 현재 시간의 차이를 의미한다. 본 연구의 목적은 인접한 교차로들이 협업을 통해 교통신호의 순서를 조정함으로써 처리하는 차량의 수를 늘리고 대기 시간을 최소화하는 것이다.

### III. 제안 알고리즘

QL은 강화학습의 일종으로 비결정적인(non-deterministic) 환경을 탐색하고 경험에 기반으로 최선의 행동을 선택한다. QL은 상태(state), 행동(action) 그리고 보상(reward)의 개념을 바탕으로 학습한다<sup>[11]</sup>. 시간이  $t$ 일 때, 환경의 상황은 상태( $s_t$ )로 정의한다. 상태에서 행동( $a_t$ )을 취했을 때, 그에 따른 보상이 주어지며 다음 상태( $s_{t+1}$ )로 넘어간다.

$$s_t \xrightarrow{a_t} s_{t+1} \quad (1)$$

$n$ 개의 상태와  $m$ 개의 행동에 대한 상태들의 집합은 식 (2) 그리고 행동들의 집합은 식 (3)과 같다. 각 상태, 행동과 보상에 대해 식 (4)와 같은  $Q$ 함수를 가진다.

$$S = \{s_0, s_1, \dots, s_n\} \quad (2)$$

$$A = \{a_0, a_1, \dots, a_m\} \quad (3)$$

$$Q: S \times A \rightarrow R \quad (4)$$

QL의 학습 값은  $Q$ 테이블(Q-table)에 저장된다. 이

때, 값은 현재 상태, 행동, 보상( $r_{t+1}$ )과 새로운 상태에 대한 값들 중에서 최댓값( $\max_a Q(S_{t+1}, a_{t+1})$ )을 학습율(learning rate,  $\eta$ )과 할인율(discount factor,  $\gamma$ )을 사용하여 구한다.

$$Q(S_t, a_t) \leftarrow Q(S_t, a_t) + \eta \cdot (r_{t+1} + \gamma \cdot \max_a Q(S_{t+1}, a_{t+1}) - Q(S_t, a_t)) \quad (5)$$

보통, QL 학습은 상태와 보상에 따라 행동을 선택하는 탐색(exploitation)을 한다. 행동을 선택할 때, 보상에 따라서 움직이는 탐색만으로 학습하는 것보다, 가끔 새로운 시도가 더 좋은 결과를 초래할 수 있다. 따라서 무작위성(randomness)을 가진 탐사(exploration)의 개념을 적용한다. 이를  $\epsilon$ -탐욕 알고리즘( $\epsilon$ -greedy selection)이라 한다.

본 연구는 다중 교차로에서의 QL을 이용한 교통 신호 제어 시스템을 제안한다. 각 교차로마다 하나의 로컬 에이전트(local agent,  $L_{agent}$ )를 가지고 있으며, 각 에이전트는 주변 인접해있는 이웃 교차로에서 해당 교차로에 들어가기 위해 대기하고 있는 차량들의 시간 정보를 받아서 독립적으로 QL을 수행한다.

제안하는 시스템은 교차로의 처리량을 최대화 하는 것을 목표로 한다. 문제의 정의는 다음과 같다.

$i$  : 교차로의 개수

$tp_i$  : 교차로  $i$ 의 처리량

$$\max \left( \sum_{k=0}^i (tp_i) \right) \quad (6)$$

이 때, 식 (7)과 같이 교차로의 개수는 로컬 에이전트의 개수와 같고, 식 (8)과 같이 한 방향의 녹색 신호의 종료와 동일한 방향의 다음 녹색 신호의 시작 사이의 간격( $t_{btw}$ )은 임계값( $\rho$ )보다 작거나 같아야 한다.

$$i = N_{L_{agent}} \quad (7)$$

$$t_{btw} \leq \rho \quad (8)$$

제안하는 QL의 상태는 해당 교차로에서 차량이 이동할 수 있는 방향의 개수로 정의하였다. 예를 들어, 그림 1에는 4-way 교차로이고, 4개가 인접해있다. 하나의 4-way 교차로는 동, 서, 남, 북, 총 4개의 방향이 있고, 각 방향마다 좌회전과 직진을 할 수 있다. 따라

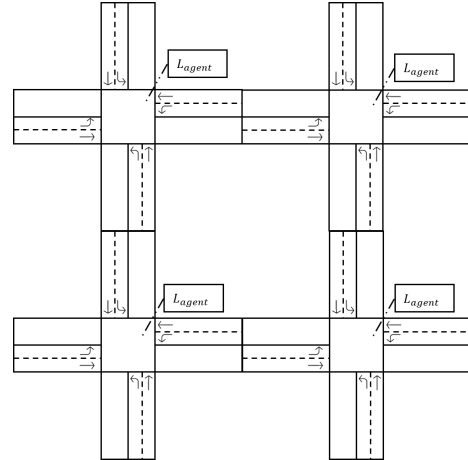


그림 1. 다중교차로의 예: 4개의 교차로  
Fig. 1. Adjacent 4 intersections.

서 하나의 4-way 교차로의 상태는 8개로 분류된다 ( $S = \{s_0, s_1, \dots, s_7\}$ ). 제안하는 QL의 행동은 교차로에서 취할 수 있는 행동으로 구성되어 있으며, 그림 2와 같이 3개의 행동 집합이 있다( $A = \{a_0, a_1, a_2\}$ ).

시간이  $t$ 일 때, 교차로에서 로컬 에이전트의 보상( $r_t^i$ )은 수식 9와 같이 처리량( $tp$ )과 인접한 교차로의 대기시간( $wt$ )로 구성되어있다. 처리량은 교차로  $i$ 에서 단위 시간 동안 처리한 차량의 개수를 의미하며, 대기시간은 교차로  $i$ 와 인접한 모든 교차로의 평균 차량의 대기시간이다. 처리량과 대기시간은 가중치 계수( $\alpha$ )로 조절하였고,  $\omega$ 는 1보다 크고  $\zeta$ 는 0과 1 사이의 값으로 정의하였다.

$$r_t^i = \alpha \times \omega_i^{tp} + (1 - \alpha) \times \sum_{k=1}^{L_{agent}} (\zeta)_i^{wt} \quad (9)$$

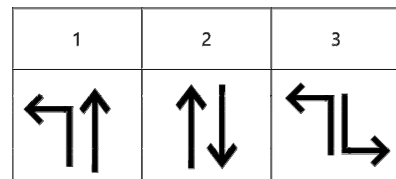


그림 2. Q-learning에서의 행동 집합  
Fig. 2. The action set in Q-learning.

#### IV. 실험 결과 및 분석

제안한 신호 제어에 대한 성능을 측정하기 위해 실

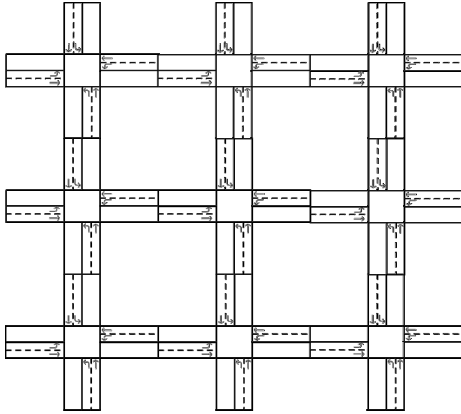


그림 3. 실험 환경에서 사용된 9개의 도로 모양  
Fig. 3. environment with 9 intersections.

험을 진행하였다. 실험 환경은 교차로의 개수에 따라 총 2가지로 분류하였다. 첫 번째는 그림 1과 같이 4개가 인접해 있고, 두 번째는 그림 3과 같이 9개가 인접해 있다. 교차로의 개수에 따라 차를 유입하는 도로의 개수가 다르다. 4개의 교차로의 경우에는 16개의 유입도로가 있고, 9개의 교차로의 경우에는 24개의 유입도로가 있다. 두 가지 환경에서 사용된 데이터는 모두 실제 데이터를 사용하였다<sup>12)</sup>. QL 학습에서 사용하는  $\gamma$ 은 0.9,  $\epsilon$ 은 0.1,  $\eta$ 은 0.01,  $\alpha$ 는 0.5,  $\omega$ 는 2 그리고  $\zeta$ 는 0.2로 설정하였다.

실험은 제안하는 시스템, 고정형 교통신호 제어 시스템과 적응형 교통신호 제어 시스템의 성능을 측정하였다. 고정형 시스템은 신호를 받는 레인의 순서와 할당되는 초록불의 시간이 항상 일정하다. 반면, 적응형 시스템은 신호를 받는 레인의 순서는 정해져있지만, 할당되는 초록불의 시간을 조절한다<sup>10)</sup>. 본 실험에서 고정형 방식은 F-TCS (fixed-time), 적응형 방식은 A-TCS (adaptive), 그리고 본 논문이 제안하는 방식은 D-TCS(distributed traffic control system)로 명칭하였다.

그림 4, 5, 6은 4개의 교차로에서의 교통량에 따라 성능 측정을 실험한 결과이다. 단위 시간당 처리할 수 있는 양을 100%라 했을 때, 유입량이 포화량보다 많은 경우에 여러 가지의 교통 문제가 발생한다. 따라서 본 실험결과는 교통량이 오버로드(overload)인 120% 부터 160%로 구성하였다.

그림 4는 4개의 교차로에서의 평균 큐 길이를 측정하였다. 평균적으로 D-TCS가 F-TCS, A-TCS보다 각각 약 50%, 55% 더 큐 길이가 짧은 것으로 나타났다. D-TCS와 A-TCS 모두 큐 길이를 최소화하고자 하였

지만, D-TCS는 인접한 교차로의 상황을 고려하였고, A-TCS는 고려하지 않은 부분에서 성능차이가 발생한 것으로 보인다. D-TCS는 교통량이 증가하여도 평균 큐 길이에 큰 차이가 없었다. 이는, 인근 교차로와의 협업이 원활하고 각 교차로에서도 적절한 신호 분배가 된다는 것을 의미한다.

그림 5는 4개의 교차로에서의 평균 대기시간을 측정한 결과이다. 평균 대기시간이란 교차로에 진입한 시점부터 차량이 교차로를 빠져나갈 때 까지 걸린 시간이다. 교통량이 120%일 때, D-TCS는 F-TCS보다 약 46% 대기시간이 더 짧다. 평균적으로 D-TCS가 F-TCS, A-TCS보다 약 45% 대기 시간이 더 짧은 것으로 나타났다.

그림 6은 4개의 교차로에서의 처리량을 측정한 결과이다. 평균적으로 D-TCS가 F-TCS, A-TCS보다 각각 약 24%, 22% 더 많은 차들을 처리하는 것으로 나타났다. 교통량이 120%일 때, F-TCS가 A-TCS보다 약 15% 처리량이 더 크다. 하지만, 교통량이 증가함에 따라, A-TCS의 처리량의 성능이 F-TCS보다 더 좋아지는 것을 볼 수 있다. 이는 4개의 교차로에서 QL

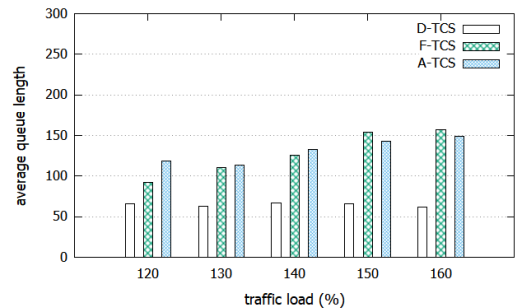


그림 4. 4개의 교차로환경에서의 평균 큐 길이  
Fig. 4. Performance comparison of queue length in 4 intersections.

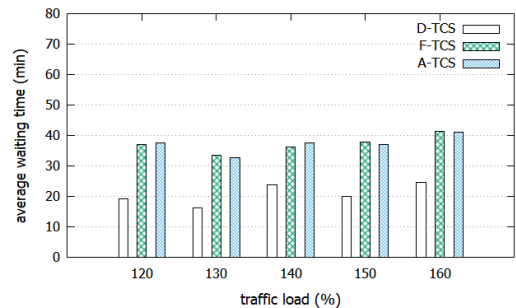


그림 5. 4개의 교차로환경에서의 평균 대기시간  
Fig. 5. Performance comparison of waiting time in 4 intersections.

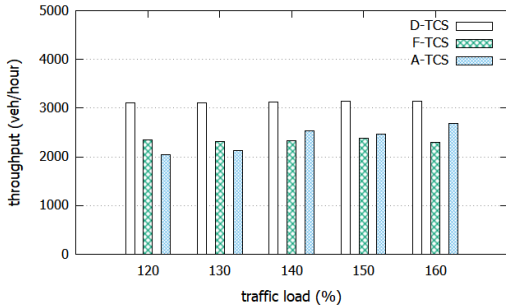


그림 6. 4개의 교차로환경에서의 처리량  
Fig. 6. Performance comparison of throughput in 4 intersections.

로 학습하는 A-TCS의 성능이 고정형 신호 제어를 사용하는 F-TCS보다 더 좋다는 것을 보여준다.

그림 7, 8, 9는 9개의 교차로에서 교통량에 따라 성능 측정을 실험한 결과이다. 그림 7은 9개의 교차로에서의 평균 큐 길이를 측정하였다. 교통량이 120%에서 160%사이인 경우, 대체적으로 D-TCS의 평균 큐 길이가 F-TCS보다 약 35%, A-TCS보다는 약 60% 더 짧은 것으로 나타났다. 교통량에 따라 초록 신호의 시간을 분배하는 A-TCS의 경우에는 교통량이 많고 인근 교차로의 교통신호 분배에 영향을 받기 때문에 큐 길이가 상대적으로 길다. D-TCS의 평균 큐 길이가 짧다는 것은 교차로에 대기하고 있는 차량이 많지 않다는 것을 의미한다. 이는 효율적으로 교통신호를 분배했음을 증명한다.

그림 8은 9개의 교차로에서 차들의 평균 대기시간을 측정한 결과이다. 교통량이 120%인 경우, D-TCS는 F-TCS와 A-TCS보다 약 32% 대기 시간이 더 짧은 것으로 나타났다. 실험 환경에는 9개의 교차로가 있고, 이들은 서로에게 영향을 끼치기 때문에 분산 교통신호 제어 시스템인 D-TCS의 성능이 다른 시스템

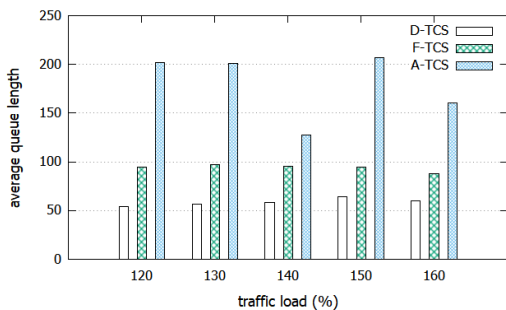


그림 7. 9개의 교차로환경에서의 평균 큐 길이  
Fig. 7. Performance comparison of queue length in 9 intersections.

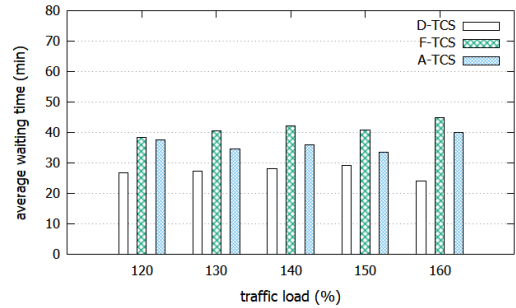


그림 8. 9개의 교차로환경에서의 평균 대기시간  
Fig. 8. Performance comparison of waiting time in 9 intersections.

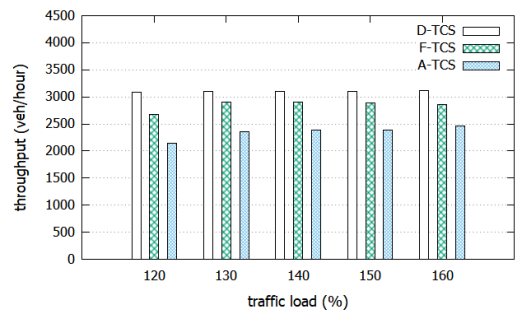


그림 9. 9개의 교차로환경에서의 처리량  
Fig. 9. Performance comparison of throughput in 9 intersections.

의 성능보다 좋게 나오는 것을 볼 수 있다.

그림 9는 9개의 교차로에서 교통량에 따른 처리량을 측정한 결과이다. 교통량이 120%일 때, D-TCS는 F-TCS보다 약 15% 차량을 더 많이 처리한다. 평균적으로 D-TCS가 A-TCS보다 약 25% 더 많은 차량을 처리하는 것으로 나타난다. 이는 D-TCS가 각 교차로마다 적절하게 교통신호를 분배했음을 의미한다.

## V. 결론

본 논문은 QL 학습을 이용한 다중 교차로에서의 교통신호 제어방법을 제안하였다. 이 연구의 목적은 인접한 교차로들과의 협업을 통해 교차로에서의 처리량을 최대화하고, 차들의 대기 시간을 최소화하는 것이다. 제안하는 시스템의 성능 평가를 위해 고정형 교통신호 방식과 적응형 교통신호 방식과 성능을 비교하였다. 결과적으로 인접한 4개와 9개의 교차로에서 D-TCS를 사용했을 때, 평균 큐 길이, 처리량, 대기시간 세 가지 측면에서의 성능이 우수하였다. 하지만, 더 큰 교차로에서의 교통신호 제어는 분산 방식만 사

용하여 교통을 통제하기에는 부족하다. 이를 해결하기 위해서 분산 방식과 중앙 방식을 혼합한 딥러닝 (deep-learning) 기반 교통신호 방법에 대한 연구가 필요하다<sup>[13-17]</sup>.

## References

- [1] F. Sivrikaya, N. Ben-sassi, X. Dang, O. Can Gorur, and C. Kuster, "Internet of smart city objects: a distributed framework for service discovery and composition," *J. IEEE Access*, vol. 7, pp. 14434-14454, Jan. 2019.
- [2] A. Zear, P. K. Singh, and Y. Singh, "Intelligent transport system: a progressive review," *Indian J. Sci. and Technol.*, vol. 9, pp. 1-8, Aug. 2016.
- [3] D. Zhao, Y. Dai, and Z. Zhang, "Computational intelligence in urban traffic signal control: a survey," *J. IEEE Trans, Systems, Man, and Cybernetics, part C (applications and reviews)*, vol. 42, no. 4, pp. 485-494, Jul. 2012.
- [4] L. Singh, S. Tripathi, and H. Arora, "Time optimization for traffic signal control using genetic algorithm," *Int. J. Recent Trends in Eng.*, vol. 2, no. 2, pp. 1-3, Nov. 2009.
- [5] Z. Yang, X. Huang, C. Du, M. Tang, and F. Yang, "Hierarchical fuzzy logic traffic controller for urban signalized intersections," in *Proc. 2008 7th World Congress on Intell. Control and Automation*, pp. 5203-5207, Chongqing, China, Jun. 2008.
- [6] I. N. Askerzada and M. Mahmood, "Control the extension time of traffic light in single junction by using fuzzy logic," *J. Electrical & Computer Sci.*, vol. 10, no. 2, pp. 52-59, 2010.
- [7] Y. K. Chin, N. Bolong, A. Kiring, S. S. Yang, and K. T. K. Teo, "Q-learning based traffic optimization in management of signal timing plan," *J. Simulation: Syst., Sci. & Technol.*, vol. 12, no. 3, pp. 29-35, Jun. 2011.
- [8] B. Abdulhai, R. Pringle, and G. J. Karakoulas, "Reinforcement learning for true adaptive traffic signal control," *J. Transport. Eng.*, vol. 129, no. 3, May 2003.
- [9] M. A. Khamis and W. Gomaa, "Enhanced multiagent multi-objective reinforcement learning for urban traffic light control," in *Proc. 2012 11th Int. Conf. Machine Learning and Appl.*, pp. 586-591, Boca Raton, USA, Jan. 2013.
- [10] Y. K. Chin, W. Y. Kow, W. L. Khong, M. K. Tan, and K. T. K. Teo, "Q-learning traffic signal optimization within multiple intersections traffic network," in *Proc. 2012 6th UKSim/AMSS Eur. Symp. Computer Modeling and Simulation*, pp. 343-348, Valetta, Malta, Jan. 2013.
- [11] B. Jang, M. Kim, G. Harerimana, and J. Kim, "Q-learning algorithms: a comprehensive classification and applications," *J. IEEE Access*, vol. 7, pp. 133653-133667, Sep. 2019.
- [12] <https://www.data.go.kr>.
- [13] D. A. Vidhate and P. Kulkarni, "Cooperative multi-agent reinforcement learning models (CMRLM) for intelligent traffic control," in *Proc. 2017 1st Int. Conf. Intell. Syst. and Inf. Management (ICISIM)*, pp. 325-331, Aurangabad, India, Oct. 2017.
- [14] C. Ma, W. Hao, A. Wang, and H. Zhao, "Developing a coordinated signal control system for urban ring road under the vehicle-infrastructure connected environment," *J. IEEE Access*, vol. 6, pp. 52471-52478, Sep. 2018.
- [15] Y. Hou, G. Wang, and Y. Zhou, "Virtual-grid based traffic control strategy with multi intersections collaboration," *J. IEEE Access*, vol. 6, pp. 40105-40119, Jul. 2018.
- [16] X. Liang, X. Du, G. Wang, and Z. Han, "A deep reinforcement learning network for traffic light cycle control," *J. IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1243-1253, Jan. 2019.
- [17] H. Nguyen, L. Kieu, T. Wen, and C. Cai, "Deep learning methods in transportation domain: a review," *J. IET Intell. Transport Syst.*, vol. 12, no. 9, pp. 998-1004, Oct. 2018.

주 현 진 (Hyunjin Joo)



2017년 2월 : 숙명여자대학교 IT  
공학과 졸업  
2017년 3월~현재 : 숙명여자대  
학교 IT공학과 석박통합과정  
<관심분야> 지능형 시스템,  
ITS, 사물인터넷  
[ORCID:0000-0003-4118-4849]

임 유 진 (Yujin Lim)



2000년 2월 : 숙명여자대학교 전  
산학과 박사  
2013년 3월 : 일본 Tohoku Uni-  
versity, Department of Info-  
rmation Sciences 박사  
2004년~2015년 : 수원대학교 정  
보미디어학과 부교수  
2016년~현재 : 숙명여자대학교 IT공학과 교수  
<관심분야> 지능형 시스템, IoT, Edge Computing  
[ORCID:0000-0002-3076-8040]