Urban Parking Scheme in Hangzhou Based on Reinforcement Learning

Mingxiu CHEN

Department of Electrical & Electronic Engineering, Faculty Office of Science and Engineering, University of Nottingham Ningbo China

Email: slymc1@nottingham.edu.cn

ABSTRACT. With the increasing number of motor vehicles in China, the traditional parking schemes have become inefficient. Traditional parking mainly relies on the driver's experience and judgment, which makes the search for a parking space difficult; at the same time it will cause traffic congestion nearby. The current parking system in Hangzhou, which includes smart payment, smart service, smart supervision and big data analysis, has largely alleviated the aforementioned difficulties. However, the existing smart parking system still has many shortcomings, such as insufficient coverage of hardware, and the software used is only an information display interface, which cannot make any judgements. This paper discusses a parking scheme based on reinforcement learning, including the application of Q-learning and DQN, in order to improve the performance of the parking system. Q-learning is the most likely method for smart parking, and DQN can also be used to improve real-time judgment. This paper compares traditional parking, smart parking, and smart parking with reinforcement learning, and lists their advantages and disadvantages respectively. The comparison shows that smart parking systems are overall more beneficial than traditional systems.

1. INTRODUCTION

With the incredible development of society and living standards, the number of vehicles in China has been increasing, and the parking problem is becoming more and more prominent. The city of Hangzhou, for example had 2.88 million motor vehicles and 740,000 parking spaces, of which about 138,000 are open to public, including highway parking spaces and 29,000 road parking garages. Meanwhile, the urban area of Hangzhou is 3,000 square kilometers, while the downtown is only 680 square kilometers, accounting for about 1/5 of the urban area of Hangzhou. Parking space compared with the developed countries in Europe and America, the proportion of urban vehicles (especially cars) and parking spaces is generally around 1:1.3, while in China's first-tier cities, it is about 1:0.8, and in small and medium-sized cities, it is about 1:0.5. However, the proportion of urban vehicles and parking spaces in Hangzhou is about 1:0.25.[1] Therefore, parking spaces in downtown are scarce, difficult to access, has many illegal parking spaces on the roadside, while there are relatively more vacant parking spaces in office buildings, large complexes, and newly built residential areas. As a result, urban parking has a very big problem, which has seriously affected daily life by making people spend a lot of time looking for parking spaces every day.

The traditional way of looking for parking spaces is mostly based on experience and can only rely on going from one parking lot to another to find vacant spaces for parking. As a result, the time cost is high and the efficiency is low; it is impossible to achieve global optimization though this way. For example, if the driver finds that there is no space in one parking lot then go to another one, the situation
is bound to add a lot of unnecessary vehicles to adjacent roads, resulting in a lot of delays and traffic jams, which greatly aggravates the traffic problem and wastes people's time.

At present, the incremental improvement scheme of traditional parking mode has been proposed. Since 2015, Hangzhou has started to build a smart parking system, which aims to cover all 740,000 parking spaces, mainly including smart payment, smart service, smart supervision and big data analysis. Smart parking system alleviates urban traffic congestion. With the help of smart parking, vehicles not only can enter and exit the parking lot quickly, but also help save a lot of social time costs. In addition to reducing traffic, license plate recognition technology with HD video technology have helped the public security department identify the deck plates, a set of HD video equipment can manage 8 to 12 parking spaces. License plate recognition rate has reached as high as 99.99%. This, coupled with video and photo storage, can provide evidence for dealing with traffic violations such as parking violations or payment evasion. There are still a few remaining problems with the implementation, such as coverage of the service and overall optimizations.

How can we avoid the path tracking error to ensure the ideal parking attitude? Reinforcement learning can be used to upgrade smart parking systems. The existing smart parking system can only display the remaining parking spaces in real time. After Q-learning is added, the program can comprehensively judge the surrounding parking spaces according to the owners' needs and driving routes, and recommend the most appropriate parking spaces. However, it will take a long time for Q-learning make judgement when state number is big. That's why DQN is used to update the algorithm. When the reinforcement learning is applied to the smart parking system as a whole, the system can make a global judgment according to the overall traffic condition, quickly obtain the optimal parking options, effectively solve the problem of finding parking spaces, and alleviate the urban traffic problems.

2. Present situation of smart parking

With the rapid development of economy and population, the supply and demand conflict between car ownership and parking space has gradually become prominent in Hangzhou, and parking difficulty has become the bottleneck to curb the progress of urbanization in Hangzhou. In order to solve this conflict, Hangzhou initiated the city brain in 2016 and embarked on a new journey of urban digital governance. Hangzhou has taken the transportation field as a breakthrough point and has started the exploration of using big data to improve the urban traffic, with significant positive results. Smart parking system includes wireless communication technology, mobile terminal technology, GPS technology, GIS technology and so on. These technologies are integrated and used in different aspects of urban parking, including collection, management, query, booking of parking spaces and vehicle navigation services. It has also realized real-time updates of parking spaces, navigation, and query. This system has significantly improved the utilization of parking space resource, the maximization of profit and the owner parking service optimization.

Smart parking is roughly composed of two components. One is traditional parking equipment, which contains a cloud parking system that can control the equipment in the parking lot to attract users through enhanced user experience. The other is Internet companies, which collect the number of vacant parking spaces and publish them on set applications. For example, Ant Financial has invested a lot smart parking area. In 2015, Alipay, a subsidiary of Ant Financial, reached a cooperation with ETCP, which is the oldest smart parking platform in China, to jointly promote smart parking. In July 2017, a pilot parking lot was set up in Shanghai Hongqiao Airport to test the "non-inductive payment". Later on, in February 2018, there is a further investment in JieShun Technology, which is a representative smart parking development company. The following figure shows a model of the smart parking system.
As for Hangzhou, in 2015, the city began to build a smart parking system, which aims to cover all 740,000 parking spaces, mainly including smart payment, smart service, smart supervision and big data analysis.

3. Smart payment
In Hangzhou, there are variety of payment methods that coexist. Along with the original support for cash and citizen card payment, it launched Alipay to scan the barcode for in-person payment in 2014, and launched Alipay to scan the QR code for self-service payment in 2015. Later on, in 2017, it took the lead in the country to launch Alipay credit non-inductive payment, and UnionPay introduced flash payment in 2018. The technical service support of non-inductive payment parking system includes smart judgment and image recognition technology, artificial intelligence technology, video parking technology, image algorithm, payment system and so on[2].

In October 2017, Hangzhou completed a pilot project of 710 parking spaces for 48 parking spots in Xihu District, which has shown good results from the operation of the pilot. It implemented credit after parking, allowing car owners to save nearly 85% of the departure time (from the traditional parking available reduced to 10 seconds, down from 70 seconds). Due to reduced payment waiting time of vehicle owners, and thus accelerate the turnover parking space, pilot area parking space turnover by the original daily average 3.8 vehicle-timeses/parking space, increased to 4.1 vehicle-timeses/parking space, which is an increase of 7.9%. The average daily revenue of a single parking space has increased from 23.12 yuan before implementation to 24.10 yuan, an increase of 4.24% and showing an increasing trend. The management range of toll collectors can be expanded from the current 20 parking spaces to 30 parking spaces, and the management efficiency can be improved by 50%.[3] By the end of 2017, more than 10,000 road parking spaces in Hangzhou have realized non-sensory payment.

4. Smart service
Smart service includes two aspects: information sharing and parking space sharing.

Information sharing: Hangzhou has built an open and compatible guidance information platform, reserved data interaction interface, and formulated data access specifications to provide technical support for the realization of parking information data sharing and the expansion of smart service parking guidance service coverage. Through "government-enterprise cooperation and resource exchange", the parking data mastered by the government can be exchanged and shared with the data collected by enterprises from social public parking garage, and the smart parking guidance service system of the city can be built.

Without government payment for this system, smart parking guidance service system has covered nearly 75% of the real-time data access of open parking spaces in urban areas, and provided users with
various query and navigation channels such as mobile apps (considerate urban management). Smart parking guidance service system guides car owners to enter a vacant parking space as soon as possible, improves the utilization rate of a parking lot, and reduces traffic congestion caused by looking for a parking space.

Parking space sharing: Hangzhou has built a parking space-sharing platform to revitalize all idle parking spaces. Hangzhou government joined Hangzhou Qianjiang New City Investment Group and Alibaba Group to build an "open, inclusive" municipal parking space sharing service platform, relying on information technology. They attempted to integrate and revitalize the parking resources of non-public parking lots, such as those belonging to enterprises, institutions, commercial buildings, residential area parking garages stock, and property owners. Release the parking space idle time to provide personalized, precision of customization service for drivers and maximum ease parking difficulty. Through the construction and operation of municipal parking space sharing platform based on credit system, the current situation of parking difficulties is effectively improved.

5. Smart supervision
Development road parking patrols regulatory mobile applications, using magnetic induction in real-time parking IoT technology acquisition data. At the same time, collect real-time parking toll collector POS machine site operation data and make comparison of these two kinds of data. Analysis of the two inconsistent data, real-time told police, make on-site supervision objectives clearer, further improve the efficiency of regulation, increased regulation deterrent.

In 2018, the social service evaluation system was added, based on the original complaint analysis to strengthen the social supervision of road parking fee service.

6. Big data analysis
This module mainly collects all kinds of data for big data analysis and realizes information sharing among government departments, which can publish "parking difficulty" index to the society, and facilitate the construction of government departments' storehouse and institutional decision-making, etc. In order to meet the needs of big data analysis and application, Hangzhou relocated the smart parking system to Hangzhou Government Administration Cloud Platform built on Aliyun "Flying Sky" in 2014, which not only reduced the IT cost by 50% compared with the original, but also reduced the cost of operation, maintenance and development of new functions, providing guarantee for mining and analyzing parking data and developing more convenient applications. In October 2018, the big data of parking system was incorporated into the urban data brain of Hangzhou, and it was extended to the design of all parking lots in the city based on the open parking garages in the main urban areas. It not only met the guidance demand of parking garage idle information, but also took into account to a series of demands such as service supervision, price monitoring, toll monitoring, traffic flow monitoring and public security.

Smart parking systems have been applied in many places in Hangzhou. For example, Hangzhou JieShun Company, along with China Unicom, and Alibaba, have built a set of new smart parking systems in demonstration area in the city, which is located in the New World Shopping Mall east of downtown. This system can realize smart parking and centralized control for the 13 parking lots of New World Shopping Mall, more than 8000 parking spaces, and it successfully connected to the city brain of Hangzhou. Through JieShun "AI Brain", the digital cockpit of urban brain can present real-time multi-dimensional information such as the flow of vehicles entering and leaving the parking lot, remaining parking spaces, temporary parking spaces, and monthly card vehicles, etc. to provide the decision-making basis for the government. In New World Shopping Mall, car owners can realize the non-inductive passage of "leave first, then pay the fee", without manual charging and opening the gate, so that the drive way is free of congestion and more smart and convenient parking services are provided for car owners. The following figure shows an integrated data platform.
To sum up, this smart parking system does improve parking efficiency from a different aspect, as it alleviates the parking problem to a great extent. However, there are still many problems. First of all, not all parking lots and parking spaces are equipped with test equipment and added into smart parking system, which mainly covers the newly built underground parking lots, and many parking spaces on the road sides are not included. In this case, the lack of input data may lead to the wrong strategy output of the converged model, which will lead to wrong guidance for people and waste more time. Additionally, even if all parking spaces are equipped with detection equipment, there is still no programs or applications to predict the driving trajectory, only real-time parking spaces can be displayed, the photo for it is shown below in figure 3, and the parking problem has not been fundamentally solved. It is easy to fall into local optimization considering the situation of one parking lot, and it is difficult to achieve global optimization. That’s why reinforcement learning should be used in urban parking scheme.

7.Q-learning for urban parking scheme
Reinforcement learning is a kind of algorithm that learns how to achieve its goal through completely random operation at the beginning, continuous trial and error, learning from mistakes, finding rules and finally learning the way to achieve its goal[4]. The emphasis is on enabling agents to learn in the environment[5]. The classification of reinforcement learning models is divided into model-free and model-based. Model-free means the policy actions the model will take is based on the type of variables are put into the model in the environment. Model-based means to understand the environment and have
expectations for the future state of the environment under specific conditions.

Q-learning is a model-free and value-based reinforcement learning algorithm, which mainly consists of four elements: agent, state, action and reward. The ultimate goal of this model is to generate an action through each time step and act on the environment to maximize the expected accumulated reward. The model of Q-learning is shown as the following,

\[
G_t = \sum_{k=0}^{T} R_{t+k+1}
\]

where,

\(G_t = \text{cumulative reward of the } t \text{ time step in the next } T \text{ time step}\)

\(R = \text{reward of every time step}\)

The formula shows that the reward for Q-learning includes the present and the future. From the time domain, the closer the distance, the more accurate it is, and the easier it is to get a reward, and vice versa; the farther the distance, the less accurate it is, and the harder it is to get a reward. The purpose of continuous model training is to obtain optimal cumulative rewards in multiple aspects while ensuring the present and future. Therefore, the discount factor is added into the formula, which makes a connection between reward and the time domain, shown below,

\[
G_t = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}
\]

where,
\(\gamma = \text{discount factor}, \text{a real value } \in [0, 1]\)

A discount factor of zero means that the agent only cares about the immediate return. The higher the discount factor, the farther rewards will propagate through time. So the value of discount factor could be used to achieve different purpose.

Q-table is used in Q-learning to represent and record the reward value of all possible actions at a certain time. The code of the algorithm is to determine whether the behavior is positive or negative, they will get a positive reward or a punishment respectively. The essence of Q-learning is a decision-making process, the results of which will be shown in the Q-table. The example of Q-table is shown below in figure 5.
In the iteration of the tested target by using reinforcement learning algorithm, agent mainly aims to find the optimal strategy and makes the environment state function \( Q(s_t, a_t) \) execute action \( a_t \) in accordance with the optimal strategy in the state of \( s_t \), so as to obtain the expectation of maximizing the cumulative reward. Moreover, since the environmental state action value function \( Q(s_t, a_t) \) can represent the expectation of maximizing cumulative reward that the agent in \( s_t \) state can obtain by taking an action behavior \( a_t \), in turn, the agent's target optimal strategy can also be obtained through the optimal value of the environmental state action function, the formula is shown below[6],

\[
Q(s_t, a_t) = E[r(s_t, a_t) + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1})]
\]

Temporal Difference (TD) method combines Monte Carlo sampling method and dynamic programming method bootstrapping (using the value function of subsequent states to estimate the current value function), which makes it applicable to model-free algorithm and is updated step by step with faster speed. The value function is calculated as follows

\[
V(s) \leftarrow V(s) + \alpha \left( R_{t+1} + \gamma V(s') - V(s) \right)
\]

Where,

\( R_{t+1} + \gamma V(s') \) is TD target
\( \delta_t = \alpha (R_{t+1} + \gamma V(s') - V(s)) \) is TD error

Base on Temporal Difference (TD) method, the Q value at time \( t + 1 \) is updating as following,

\[
Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[ R(s_t, a_t) + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t) \right]
\]

Where,

\( \alpha \in (0, 1) \) is the learning rate, which can balance the new and old information.

For example, a learning rate of zero means that the agent only uses old information, no further Q value is update required. A learning rate of one means that the agent only uses new information, Q value is completely updated with the latest information.

As for urban parking scheme, three elements (state, action and reward) are defined.

Action: at a certain time step \( t \), the action is defined to execute the parking process of the vehicle, which contains park and no-park. If the agent decides to execute the park process, the vehicle would park at that parking area at time \( t+1 \). On the contrary, if the agent decides to execute the no-park process, the vehicle would keep searching for the next parking area at time \( t+1 \).

State: state contains four different aspects, which are number of remaining spaces \( n_s \), size of parking lot \( p \), cost of parking \( c \), and real time distance \( d \) between the vehicle and that parking lot, all normalized between 0 and 1. The input matrix is used here to bring the four aspects together, an example is shown below,

|   | a1 | a2 |
|---|----|----|
| s1 | -5 | 10 |
| s2 | -2 |  3 |
| s3 |  6 | -7 |
$S = [n_s, p, c, d]$

Reward: different states represent different values in reward at each time step $t$; the sum of every kind of reward becomes the total reward which will show in the Q-table.

$$R = R_n + R_p + R_s + R_d$$

As for traditional urban parking scheme, people tend to make a judgment about where to park based on experience. In fact, it is difficult for human brains to make a judgment based on experience when global optimization is achieved and multiple factors such as parking time, parking price and moving distance are taken into account. As a result, reinforcement learning is used to solve the problem. With Q-learning, the computer can give the optimal solution based on the surrounding environment.

In real life, for complex urban parking scheme application scenarios, each parking lot’s parking spaces needs to be updated in real time, changes near the parking lot’s traffic situation has to be included in the consideration. Therefore, there are too many state number of $Q(s_t, a_t)$ behavior, the problem will appear as space is too large, so using Q-learning machine learning alone will lead to a long calculation time, which usually causes delayed decision; amount of memory and data also are a problem.

8. Value Function Approximation and DQN

In order to solve the problem of large state space, Value Function Approximation is applied. Linear or nonlinear functions replace the Q-table to represent $Q(s_t, a_t)$. The optimal objective function is generally the mean square error objective function, and the optimization strategy is generally the gradient descent optimization.[7] The approximate function is shown as following,

$$\hat{v}(s, \omega) \approx v_t(s) \text{ or } \hat{q}(s, a, \omega) \approx q_t(s, a)$$

Where,

$\omega =$ The weight of the neural network

Weight is a parameter in neural network, which converts input data in the hidden layer of neural network. A neural network is a series of nodes, or neurons. Each node has a set of input, weight, and deviation values. When an input is sent to a node, it is multiplied by a weight, and the resulting output is either observed or transmitted to the next layer of the neural network. Usually, the weights of neural networks are included in the hidden layer of neural networks[8].

In order to fit the appropriate function and calculate the weight, the supervised learning algorithm in the machine learning algorithm is applied, which mainly include references to regression algorithms, such as linear regression, decision tree, neural network, etc. The input state is extracted with features as input, and the value function is calculated through MC/TD as output. Then the function parameters are trained until convergence.

Q-learning algorithm is a kind of off-policy learning algorithm, that is to say, it has two strategies: target strategy $\pi$ and action strategy $\mu$. When the Q-value of a state behavior pair is updated, it is not the Q-value of the next state behavior pair that currently follows the policy, but the Q-value of the next state behavior pair that is generated by the target policy that is being evaluated.

As a combination of Q-learning and the neural network, DQN (Deep Q-Network) is considered to be used in urban parking scheme. The feature of off-policy is also extended in DQN. The difference is that the Q used to calculate target and predicted value in Q-learning is the same Q, that is, the same neural network is used. One problem with this is that every time the neural network is updated, the target will also be updated, which will easily lead to non-convergence of parameters. In supervised learning, labels are fixed and do not change with the update of parameters.

Therefore, DQN introduces a Target Q network on the basis of the original Q network, that is, the network used to calculate the Target. It is the same as the Q network structure, with the same initial weights, except that the Q network is updated every iteration, whereas the Target Q network is updated every once in a while.

Compared with Q-learning, DQN has some improvement. For example, DQN uses convolutional neural network to approach the behavior value function. Additionally, it uses Target Q Network to update the target. Lastly, it uses experience replay.
In reinforcement learning, the observed data obtained are in great order. It will be a problem to update the parameters of the neural network with such data. In supervised learning, the data are independent of each other. Therefore, experience playback is used in DQN, that is, a Memory is used to store experienced data, and a part of the data is extracted from Memory for updating each parameter, so as to break the correlation between data.

After replacing Q-learning with DQN, the problem of huge state space in urban parking scheme can be greatly improved. The model of applying DQN in urban parking scheme is shown in the figure below,

![Figure 6. DQN model in urban parking scheme](image)

9. The application prospect of reinforcement learning in smart parking and transportation

The current level of hardware and software is not enough to change the traditional traffic industry; part of reinforcement learning can only be applied to specific scenarios such as those mentioned in this paper. Smart parking planning, real-time road data (such as access to relevant data traffic from Google Maps) can make the decision more accurate, and achieve global optimization easier. At the same time however, it also makes data processing more complex and difficult, and there is still great room for improvement in the optimization of reinforcement learning algorithm.

With the accumulation of historical parking spaces data and traffic data in cities, as well as the development of artificial intelligence and big data technology, urban parking and even the entire traffic system will develop towards data-oriented intelligence, such as the application of artificial intelligence in traffic lights, traffic flow prediction, traffic congestion prediction and so on. The whole city’s traffic can be uniformly distributed, from route planning to daily parking. When the smart technology of traffic decision-making is fully matured, the technology of fully autonomous driving vehicle will also become possible.

Under such circumstances, the widespread application of 5G communication and the development of computer hardware also appear particularly important. Smart data, reinforcement learning and real-time decision-making are based on a large number of data transmission and computer simulation, and if the network speed and hardware equipment are not up to the standard, it will inevitably affect the smart process of the traffic system.

REFERENCES

[1] X. Shuo. "The smart parking industry is expected to reach 10 billion yuan." http://www.its114.com/html/news/survey/2018_12_99621.html (accessed 9.10, 2020).
[2] "The technology behind smart payment." http://www.dimeng.vip/news-baike/41580.html (accessed 8.25, 2020).
[3] H. R. P. F. S. Center. "Fine management wisdom parking hot spot is coming." https://www.sohu.com/a/299898308_776618 (accessed 9.13, 2020).
[4] R. Sutton and A. Barto, Reinforcement Learning: An Introduction. MIT Press, 1998.
[5] C. J. C. H. Watkins and P. Dayan, "Q-learning," Machine Learning, vol. 8, no. 3-4, pp. 279-292, 1992.

[6] D. Pandey and P. Pandey, "Approximate Q-Learning: An Introduction," in 2010 Second International Conference on Machine Learning and Computing, 9-11 Feb. 2010 2010, pp. 317-320, doi: 10.1109/ICMLC.2010.38.

[7] X. Li, H. Liu, and X. Wang, "Solve the inverted pendulum problem base on DQN algorithm," in 2019 Chinese Control And Decision Conference (CCDC), 3-5 June 2019 2019, pp. 5115-5120, doi: 10.1109/CCDC.2019.8833168.

[8] "Weight (Artificial Neural Network)." https://deepai.org/machine-learning-glossary-and-terms/weight-artificial-neural-network#:~:text=Weight%20is%20the%20parameter%20within%20a%20neural%20network%2C%20set%20of%20inputs%2C%20weight%20and%20bias%20value (accessed 8.24, 2020).