Deep Reinforcement Learning based Tourism Experience Path Finding

[1]Kyung-Hee Park, [2]*Juntae Kim

Abstract

In this paper, we introduce a reinforcement learning-based algorithm for personalized tourist path recommendations. The algorithm employs a reinforcement learning agent to explore tourist regions and identify optimal paths that are expected to enhance tourism experiences. The concept of tourism experience is defined through points of interest (POI) located along tourist paths within the tourist area. These metrics are quantified through aggregated evaluation scores derived from reviews submitted by past visitors. In the experimental setup, the foundational learning model used to find tour paths is the Deep Q-Network (DQN). Despite the limited availability of historical tourist behavior data, the agent adeptly learns travel paths by incorporating preference scores of tourist POIs and spatial information of the travel area.

Keywords: *Reinforcement Learning, Path Finding, Tour Planning, Smart Tourism, Digital twin*

I. Introduction

The rapid advancement of information technology and the increasing popularity of Free Independent Travel (FIT) [1] have led modern tourists to rely on smart devices with constant internet connectivity during their tours [2]. Consequently, there is a growing demand for services that utilize big data and machine learning to understand tourists' preferences and offer personalized tour recommendations [3].

However, due to the limited availability of data on tourist behavior [4], machine learning often lacks access to substantial information. The majority of existing tourism data is related to statistical records of tourism product sales and content shared on social media platforms, which do not provide a deep understanding of tourists' movements within tourist destinations.

To address this challenge, we introduce the 'Tour Path Finding Algorithm based on Reinforcement Learning,' which makes use of both Point of Interest (POI) data and geographical spatial information from a digital twin. Even without historical tourism data for the specific tourism area, this algorithm efficiently leverages real-time Point of Interest (POI) data and spatial information for learning, facilitated by the digital twin. Consequently, it empowers tourists to enhance their tourism experiences by identifying optimal tour paths.

II. Machine learning and path finding problems

The proliferation of mobile internet devices and the availability of location-based services have made it increasingly convenient to track the travel paths of tourists [2]. Ongoing research focuses on leveraging travel path data in combination with tourism-related information, including internet and social media data, for the purpose of tourism planning and travel recommendation services.

[1] *Ph. D Candidate in Dept. of Computer Engineering, Dongguk Univ. (kyunghee.park@gmail.com)*

[2*] Corresponding Author *Univ. of Dongguk, Dept. of Computer Engineering, Professor (jkim@dongguk.edu)*

Received: Oct. 26, 2023, Revised: Dec. 15, 2023, Accepted: Dec. 21, 2023

Lee [5] introduced a system that harnesses Recurrent Neural Networks (RNN) to analyze the sequence of tourists' visits to Points of Interest (POIs) and infer the significance of this sequence to provide comprehensive travel routes. This method learns from the historical order in which past tourists visited POIs and offers recommendations for the order in which users should visit POIs.

Chang [6] proposed a system that identifies Regions of Interest (ROIs) encompassing POIs visited by package tourists in the past and recommends package tourism products based on the visit order of these ROIs. This system utilizes reinforcement learning to understand travel-visit orders within ROIs, incorporating demographic data.

Geng [7] demonstrated that reinforcement learning can be applied to address the problem of finding the optimal walking path, making use of real-time population density data along the available paths.

Park [8] illustrated the effectiveness of reinforcement learning in solving the optimal path search problem, taking into account pedestrian movement characteristics.

These approaches contribute to enhancing tourism planning and improving the travel experience for tourists by utilizing data-driven and machine learning techniques. However, previous research on tourism recommendation methods has primarily focused on suggesting tourist Points of Interest (POIs) considered attractive to tourists. The challenge of determining the most advantageous path for tourists to reach these recommended POIs has not been fully addressed.

Tourists often venture into unfamiliar territories in search of undiscovered POIs, adding to the sense of novelty and adventure in their travels. While digital maps, primarily accessed by individual tourists through smartphones, typically provide path guidance based on factors like the shortest travel time or distance, these approaches often fall short in capturing the true essence of tourism, resulting in missed opportunities for more enriching tourism experiences. [9]

Considering tourism as a purposeful endeavor to create lasting impressions, our objective is to explore travel paths that facilitate the accumulation of diverse and memorable tourism experiences. Leveraging reinforcement learning, we aim to identify optimal paths for tourists that offer maximum engagement with a wide range of enriching encounters within a designated region. Reinforcement learning (RL) agents adaptively learn distinct trajectories within the tourist locale, guided by a reward function that evaluates the tourism experience along these paths.

As a result, the introduced algorithm derives insights from up-to-the-minute data that reflect fluctuations in tourism trends and changes within tourist destinations, sourced directly from the digital twin. This approach distinguishes itself from methodologies dependent on extensive historical tourism behavior data.

III. Reinforcement Learning-Based Tourism Experience Path Finding

3.1. Proposed Algorithm

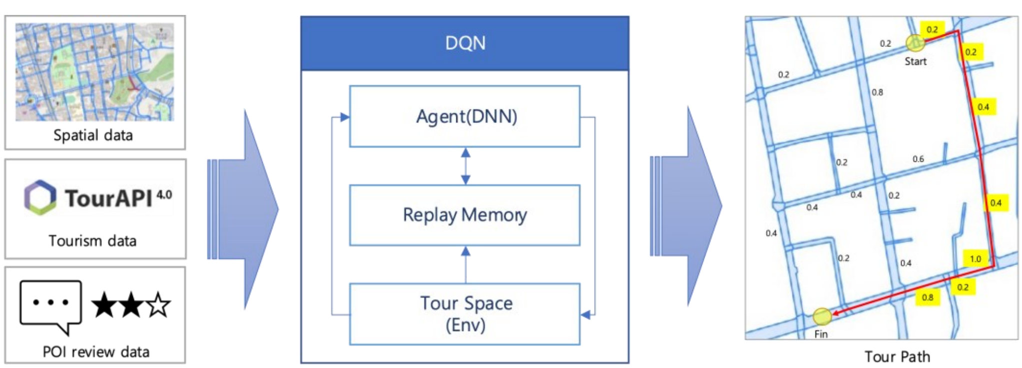

Fig. 1 Reinforcement Learning based Tourism Path Finding

As previously mentioned, the tourism sector is currently undergoing a dynamic integration of Information Communication Technologies. Prominent tourist destinations are exploring the adoption of digital twin technologies to enhance both the tourism experience and the overall industry [10]. Tailored digital twins for tourism have the capacity to provide real-time insights into the tourism environment, offering information about POI(Point of Interest), including details like operating hours and occupancy status, as well as spatial data within the region. The feedback and evaluations contributed by tourists through their reviews are regarded as invaluable datasets, and digital twins acquire this data in real-time. Figure 1 illustrates the experimental framework we have developed, with the input data depicted on the left, demonstrating the existing information utilized by experimental digital twin systems tailored for tourism.

Reinforcement learning, a paradigm in which an agent learns actions to take within a given environment, focuses on acquiring optimal action policies [11]. The outcomes of these actions should result in the highest cumulative rewards, and for this purpose, a measurable reward function is indispensable. In our tourism experience path planning, the reinforcement learning environment is constructed using tourism data from 3D maps [12] (depicting spatial information) and TourAPI [13] data (containing details about tourist Points of Interest or POIs) sourced from the experimental digital twin. The reward function incorporates a tourism experience index derived from the evaluations provided by previous visitors, which include the number of reviews and ratings of the visited POIs. As a result, the reward function finds an equilibrium between travel costs (such as time or distance to travel) and the quantity of the tourism experience.

Deep Q-Networks (DQN) [14] is a deep learning model designed to directly learn control policies from complex sensory inputs. In this study, we employ DQN to address the challenge of identifying optimal paths to enhance the tourism experience. DQN includes a decoupled target neural network (Θ-) responsible for computing target values. This architecture resolves the issue of simultaneous shifts in Q network parameters during the learning process, thus improving learning performance.

DQN, originally developed to address video game challenges, has demonstrated its ability to effectively identify and analyze game elements within intricate state spaces, often referred to as screen space. In the context of the tourism path finding problem, the challenge lies in discerning spaces in which various Points of Interest (POIs) are distributed across an extensive tourist area. The DQN model has proven to be effective even in tourist route spaces that bear a resemblance to the complexity found in video game environments.

DQN based tourism path finding algorithm used in this study is as follows.

```
Algorithm : DQN based tourism path exploration
Initialize Replay Buffer B
Initialize Q network with random
for episode e=1 to max episode do
    Initialize Env, get initial position s1
    For time_step t=1 to max time_step do
        Chose Action a(t) using epsilon-greedy method with a fixed parameter ε
        Move to the direction indicated by A(t), get reward r(t) from Env and observe next
position s(t+1)
        Collect experience and save e(t) to B
        Get sample batch experience (s(j), a(j), r(j), s(j+1)) from B
        If episode terminates then : Calculate y(j) with formula 8
        Else  y(j) = r(t) end if
        Perform gradient descent and update Q-network
        Update current state
        Every k step set T- = T
    End for
End for
```

3.2. Experiment

We defined the state, action, and reward of reinforcement learning for tourism experience path finding as shown in Table 1 below.

Table 1. State, Action, Reward definition for Tourism Experience Path Finding

Elements	Definition
State	$S = V = \{v(x, y) \mid v \in V, x = Longitude, y = Latitude\}$
Action	$A = \{N, E, S, W\}$
Reward	$r_t = -c_{ij} + \omega_d \cos\theta + \omega_p \phi(t) + 1(s_t = s_{end})$

(1) State: The state is roads and its direction of a region that agent can move from the current location to other.

(2) Action: The action that agent can take at any location v is to select an edge that can move to another intersection point v' connected to v. All possible directions can be expressed as an azimuth divided from 0 to 360 degrees clockwise from the current position due north. This research employed multiples of 4 as movement directions to enhance learning efficiency, including options for 4, 8, and 16 directions. The action can be further tailored based on the existing road directions. However, in most modern planned urban spaces, freedom of movement can be efficiently represented with just four primary directions (actions). In this paper, for the sake of learning efficiency, the available moving directions (Action) were constrained to four options: due north (N, 0), due east (E, 90), due south (S, 180), and due west (W, 270). The direction of all roads was approximated to the nearest of these four cardinal directions.

(3) Reward: The reward function consists of the cost of a movement and acquired tourism experience from the movement. c_{ij}, the first term of reward function, is the cost (in time or distance) of moving from v_i to v_j. $\omega_d \cos\theta$, second term, is the moving direction component that guides agent to the expected destination in the vast state space at last. The third term $\phi(t)$ is the tourism experience value which agent can gets from the moving, and finally, when the agent successfully arrives at the destination, 1 is additionally given as a success reward.

We have developed an experimental digital twin of a specific tourism region in Korea, covering an area of approximately 90,000 square meters, which serves as the reinforcement learning environment. This digital twin incorporates 3D maps and tourism Points of Interest (POI) data collected from online sources [15]. Figure 2 illustrates the layout of a portion of the digital twin, where each numerical value (ranging from 0 to 1.0) represents the normalized tourism experience calculated for a particular segment of the path. Throughout the experiment, the tourism experience metric was computed based on the number of encountered POIs along the given path and their review scores from previous visitors.

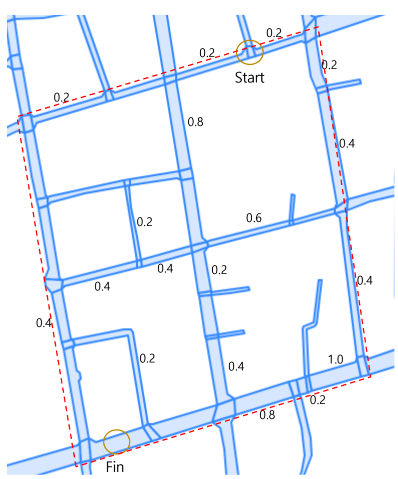

Fig. 2 Experimental Environment (Tour Space)

The red dotted line of the map is borderline of tour area allowed that RL agent can move while its learning.

In this study, DQN is configured with the Q-network with 50 neurons in each of the two hidden layers, $\varepsilon = 0.05$, $\alpha = 1.0e\text{-}4$, $\gamma = 0.95$, and ReLu was applied for the activation function. These hyperparameter values are the result of finding a balance between performance and efficiency through repeated experiments.

The tourism experience (ϕ) associated with individual Points of Interest (POIs) was set up to provide the agent with a tourism experience along the connected road segments. The tourism experience that each POI conveys to the agent through its adjacent road segment is calculated using equation (1). POIt represents the tourism experience value associated with POI t. In this experiment, the tourism experience value of each POI is computed based on the number of reviews and the review scores provided by previous tourists.

$$\phi_{POI}(t) = \begin{cases} \dfrac{POI_t - POI_{min}}{POI_{max} - POI_{min}} & (visit\ count = 1) \\ 0 & (visit\ count > 1) \end{cases} \qquad (1)$$

The value of ϕ was considered valid solely for the initial visit within each episode. This approach was implemented to deter the agent from repeatedly traveling along the same road segments within a single episode, which could artificially inflate its tourism experience values.

During the training process, we conducted a total 10 runs, each consisting of 100 episodes.

3.3. Experiment Results

Agents explore optimal travel paths to maximize tourism rewards while traveling from the origin to the destination. This tourism experience can be compared with other path search algorithms by evaluating the tourism experience (Φ) per unit distance of the path (Φ/km).

(1) φ : 1.6, Distance : 0.58km (2) φ : 3.2, Distance : 0.78km (3) φ : 3.8, Distance : 1.29km

Fig. 3 Tourism experience paths founded by DQN

Figure 3 illustrates the agent's moving trajectory overlaid onto the map, along with the corresponding tourism experience attained through each action. The trajectory labeled as (1) in Figure 1 corresponds to the outcome of the A* algorithm, a conventional approach for finding the shortest path. In contrast, trajectory (2) in Figure 1 yielded the highest recorded tourism experience value, effectively doubling the tourism experience compared to trajectory (1), albeit at the cost of covering an additional 34% distance. The tourism experiences collected as a result of the experiment and the distance traveled are shown in Table 2.

Trajectory (3) in Figure 3 is characterized by the agent retracing its steps to a previously traversed segment, subsequently seeking a novel path to access more tourism experience (depicted within the orange shaded box). This behavior closely mirrors the proactive tendencies of a human tourist exploring new avenues while on a tour.

Table 2. DQN Tourism Experience Path Finding Results

Path	Φ/km	Moving Distance (km)	Tourism Experience (φ)
A* Algorithm	2.76	0.58	1.6
DQN (Max Reward Path)	**4.1**	0.78	3.2
DQN (Max Experience Path)	2.95	1.29	3.8

DQN began to consistently perform well after approximately 20 training sessions. As shown in the left graph of Figure 2, despite fluctuations in reward values, DQN demonstrates a noticeable trend toward converging to an optimal solution over the learning iterations. The right graph of Figure 1 illustrates irregular oscillations in cumulative reward values as episodes progress, a characteristic phenomenon in reinforcement learning processes.

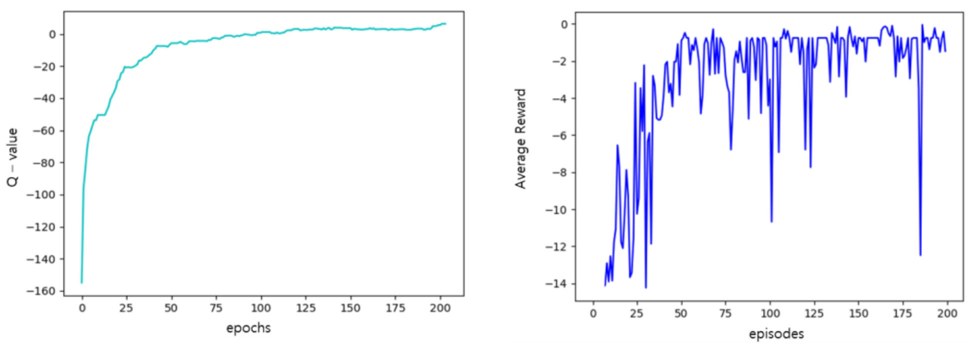

Fig. 4 DQN model training graph (L: Q-Value, R: Average Reward)

IV. Conclusion

Reinforcement learning agents that learn from both spatial information and Points of Interest (POI) data have demonstrated proficiency in identifying optimal tourist paths tailored to travel objectives. The algorithm introduced in this study offers several advantages over previous research.

The suggested algorithm not only discovers recommendable POIs but also recommends paths aligned with the intended purposes of the tour. This achievement stems from the agent's ability to learn in real-time about the dynamic conditions within the tourist region with the assistance of digital twin technology such as POIs operating status and changes in tourism trends.

Even in environments where access to historical data of past tourist behaviors is limited, the tourism experience path finding algorithm remains effective in identifying and suggesting optimal tourist paths. This capability is realized by learning POI data along with the spatial characteristics of the respective tourist area from the digital twin or the internet. The proposed algorithm demonstrates practical utility even in tourist areas with newly established attractions or in situations where accessible historical data for learning is scarce.

V. Acknowledgments

This research was funded by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport, grant number 1615012545

VI. References

[1] Parr, Diana K. Free independent travellers: The unknown tourists. Diss. Lincoln College, University of Canterbury, 1989.

[2] Wang, D., S. Park, and D. Fesenmaier, "The role of Smartphones in Mediating the Touristic Experience", Journal of Travel Research, Vol.51, No.4, 2012.

[3] Jongwook Lee and Woontack Woo. Current Status and Prospect of Smart Tourism Using 3D Map Linked Digital Twin. The Journal of The Korean Institute of Communication Sciences,36(10),55-62. (2019).

[4] Popescu, Gabriela, et al. "Current problems regarding the research of consumer behavior in tourism." Lucrări Ştiinţifice Management Agricol 18.1 (2016): 269.

[5] HeeJun Lee and Choong Kwon Lee. Sequence-Based Travel Route Recommendation Systems Using Deep Learning - A Case of Jeju Island -. Smart Media Journal,9(1),45-50. (2020).

[6] Chang, Jui-Hung, et al. "Travel Package Recommendation based on Reinforcement Learning and Trip Guaranteed Prediction." Journal of Internet Technology 22.6: 1359-1373 (2021).

[7] Geng, Yuanzhe, et al, "Deep Reinforcement Learning Based Dynamic Route Planning for Minimizing Travel Time." 2021 IEEE International Conference on Communications Workshops (ICC Workshops), IEEE, (2021).

[8] Park, Hyoshin John,"A Personalized Trip Planner For Vulnerable Road Users." Center for Advanced Transportation Mobility. 8. (2019).

[9] Majid Alivand, et al, Analyzing how travelers choose scenic routes using route choice models, Computers, Environment and Urban Systems, Volume 50, 2015, Pages 41-52, ISSN 0198-9715

[10] Ssin, S., Suh, M., Lee, J., Jung, T., Woo, W. (2021). Science Tour and Business Model Using Digital Twin-Based Augmented Reality. In: tom Dieck, M.C., Jung, T.H., Loureiro, S.M.C. (eds) Augmented Reality and Virtual Reality. Progress in IS. Springer, Cham.

[11] Richard Sutton, "Reinforcement Learning: An Introduction", 2017

[12] Digital Twin National Land 2023, V-World, https://map.vworld.kr/map/ws3dmap.do?mode=MAPW201

[13] Korea Tourism Organization 2023, Tour-API, https://api.visitkorea.or.kr/#/useUtilExercises?utilName=&requestPage=1

[14] Volodymyr Mnih, et al, "Playing Atari with Deep Reinforcement Learning", arXiv preprint arXiv:1312.5602. (2013)

[15] Kyung-Hee Park and Choel-soon Kim, Smart Tourism Contents based on Augmented Reality combined with Digital Twin, ICT Platform Society 2022 Conference , PTL Volume 9-1, Pages 41-47, ISSN 2288-8195

Authors

Park, Kyung-Hee

2008.7 ~ : Clayworkd Inc., CEO
2021. 8. : Ph.D Candidate in Dept. Of Computer Engineering, Dongguk University

Research Interests : AI, Metaverse, Digital Twin

Kim, Juntae

1995 ~ : Professor in Dept. Of Computer Science & Engineering, Dongguk University, Korea
1994~1995 : Southern Methodist University, USA. Postdoctoral Research Associate.

Research Interests : AI, Machine Learning, Meta Leraning