Implementation of connection scan algorithm in tourism intermodal transportation journey planner: a case study

Vertic Eridani Budi Darmawan¹*, Yuh Wen Chen²
¹Department of Industrial Engineering, Universitas Negeri Malang, Jalan Semarang No. 5, Malang 65145, Indonesia
²Institute of Industrial Engineering and Management, Da Yeh University, No. 168 University Rd. Dacun, Changhua 51591, Taiwan

ABSTRACT

Accessibility to tourist destinations is an important component in a tourism system, especially for natural tourist destinations located in suburban areas. Good linkage of travel information and physical connections with local transportation services for intercity travel can facilitate more people to travel and promote national tourism destinations. This research takes the popular national tourism destinations and their public transportation service in Taiwan as a research object due to the unavailability of integrated public transport information service. Free Independent Travelers (FIT) demand is growing. This research aims to integrate intermodal public transportation information to support FIT by proposing a seamless way journey planner. In this scenario, the journey planner requires timetable data as input. The Connection Scan Algorithm is used to find the earliest arrival time routes at their destinations. This journey planner is built in PHP language and can complement the official tourism travel information website by Tourism Bureau, MOTC. Hence, the FIT could get the quickest routes to reach the destinations without compiling the public transportation information provided independently.

Keywords:
Journey planner
Intermodal transportation
Tourism
Connection scan algorithm
Free independent travelers

1. INTRODUCTION

The rapid growth of tourism lies in the high trend towards globalization, and the world continues to be more connected. The positive effects of globalization on tourism supply and demand are increasing economic prosperity, creating employment opportunities, spreading the usage of Information Communication and Technology (ICT), developing new tourism markets, and changing social structure [1]. Meanwhile, it also leads to an adverse impact on the deterioration of the natural, social, and cultural environment [1]–[3]. Tourism exacts a consequence of change, and it’s required to be well planned and managed. The successful tourism development depends on maintaining a delicate balance, which is to get a full benefit to outweigh the cost and lessen the detrimental impacts as a part of the change. It should be done in a sustainable way of economic development [4].

Tourism could not possibly occur without transportation, and it’s the most profound in the

*Corresponding Author
Vertic Eridani Budi Darmawan
E-mail: vertic.eridani.ft@um.ac.id

This is an open-access article under the CC-BY-NC-SA license.
growth of global interconnectedness [5]. As a key element of the tourist experience, transportation provides the main connection between origin and destination tourism area and relieves tourists' mobility [5, 6]. The accessibility of transportation between origin and destination, will shape how tourist behaviors, especially whether they use public or private transportation [6]. Concurrently, providing the information to motivate tourists to take public transportation at the tourism destination area, is one of 21 Tourism Agenda in Sustainable Tourism [7].

In Taiwan's case, most international tourists would rather be Free Independent Travelers (FIT) [8]. Most of them also searched about Taiwan’s tourism information through the Internet before they arrived in Taiwan. After they came, the transportation information (60%) is the most information they were searching for, then followed by scenic spots (49%) and restaurant information (49%) [8]. The awareness of travel information-seeking behaviors could be as a prospecting guide for the government agencies and businesses to decide an effective transportation service delivery for the tourists [9].

Given the increasing number of tourists, some initiative has been made to publicize public transportation used by tourists in Taiwan, especially in the urban area, such as Fun Travel in Taipei (FIT). Traveling across Taiwan (intercity travel) could be convenient too because Taiwan’s Government also offers innovative and intimate travel service apps called “Tour Taiwan” and “VZTaiwan”. Both apps cover tourism destinations all over Taiwan. They can also suggest a planned trip depends on tourist preference as the day's limit, the theme of the holiday, and itinerary pace.

However, these services still lack reliable connection information to the public mode traffic, and the service apps build independently with the main tourism website. Meanwhile, the advanced traveler information service is needed to help FIT decide to reach the tourism destination with absolute certainty of transportation information [10]. The intercity traveling in Taiwan has depended much more on public transportation service [11].

Providing tourism information services, especially transportation information, is critical in the globalization of the tourism industry. The tourism service shapes new tourism markets that tourists become more independent and freer to travel (FIT). The accessibility of transportation between origin and destination will shape how tourist behaviors, especially whether using public or private transportation [6] and the cooperation of transportation modes, lead to improved transportation experience and efficiency of the transportation systems [12]. Taiwan has a high-quality public transportation system [13]. The development of an intermodal passenger transportation system continues to be improved [14]. It provides an excellent context for integrating an existing public transportation system to the tourism destination by developing the journey planner.

The behavior of persons using public transportation is different from traveling in private mode. The persons are only able to depart and arrive at the vehicle at specific stations within the networks. Consequently, they need an alternative to find routes from their origin to the destinations. The availability of routes in the public transportation network depends on the timetable data. It raises the timetable information problems by finding the earliest arrival routes at their destinations within the network [15].

Over the years, many research types contributed to routing algorithms to solve this problem [16]. Two major approaches have been proposed to solve the timetable-based shortest path problem: graph-based and non-graph approaches. The graph-based process is extended Dijkstra’s algorithms with the two common methods called the time-dependent and time-expanded [15]. This approach could not give convincing results or lead to high query times when applied to public transportation networks with a different structure than the road network [17].

The first non-graph-based approach was RAPTOR (Round-based Public Transit Routing) [18]. RAPTOR utilized the essential public transit network timetable elements, and it has no need preprocessing process and works in fully dynamic scenarios. Currently, RAPTOR’s algorithm is behind software such as Open Trip Planner [16].

No longer after RAPTOR algorithm was invented, the Connection Scan Algorithm (CSA) was introduced in 2013 [19]. CSA uses an approach of modeling the timetable data as a directed acyclic graph or elementary connection. Topologically, the primal connection is sorted by departure time, and the shortest path algorithm only needs to scan all the connections of the timetable in a single array once. Currently, CSA is behind the mobiTopp, a modular agent-based
Travel Demand [20] and used in open data initiatives for the public transportation route planning ecosystem [21].

The algorithms are often built to answer slightly different questions and are specially implemented on top of other data models. Instead of developing a new algorithm, this research is looking for trouble-free implementation with the smallest building block needed to fit our needs. CSA is slightly more efficient than RAPTOR, but it is substantially straightforward [16]. Hence, in this research, CSA was selected as the proven fastest algorithm to solve our problem in public transportation route planning than traditional route planning solutions based on variants Dijkstra’s algorithm [16], [20], [22].

The purpose of this study is to integrate intermodal public transportation information to support FIT by proposing a seamless way journey planner that has not been addressed yet. In this scenario, the journey planner requires timetable data as input. The Connection Scan Algorithm is used to find the earliest arrival time routes at their destinations.

2. RESEARCH METHODS

2.1 Connection Scan Algorithm

The Connection Scan Algorithm (CSA) was firstly introduced in 2013 [19] and well-proven as an efficient public transportation route planning [22]. The elementary connection is a fundamental unit to compute the earliest arrival time [22]. It is the basic building block of the timetable, and CSA efficiently arranges it into a single array. CSA computes the network by scans the connection array linearly to get the journey plan. The basic problem idea that CSA offer is the earliest arrival time and extends to handle multi-criteria profile queries such as the earliest arrival time and the minimum number of transfers [22].

In CSA, the timetable for public transportation could define in a mathematical formulation as a tuple:

\[
(S, C, T, F)
\]

where
- \(S\) is a set of Stations
- \(C\) is a set of Connections
- \(T\) is a set of Trips
- \(F\) is a set of Footpaths.

**Definition 2.1.**

A station is a place where the passenger can stand and halt to leave and enter the vehicle.

\[
S = (s_1, s_2, s_3, \ldots, s_n)
\]  \hspace{1cm} (2)

Time, when the vehicle arrives and departs at a station, is well known. The vehicle arrives at each station \((s)\) at arrival time \(t_{arr}\) and departs at the departure time \(t_{dep}\). The times are fully described in the timetable.

The foundational element in the timetable is a connection \((c)\). It represents a vehicle that drives between one stop to another stop without any intermediate halt [22].

**Definition 2.2.**

A connection is a tuple

\[
c = (s_{dep}, s_{arr}, t_{dep}, t_{arr}, t_r)
\]  \hspace{1cm} (3)

where:
- \(s_{dep}\) is the departure station
- \(s_{arr}\) is the arrival station
- \(t_{dep}\) is the departure time
- \(t_{arr}\) is the arrival time
- \(t_r\) is a trip id

with \(s_{dep} \neq s_{arr}\) and \(t_{dep} \leq t_{arr}\).

By definition 2.2, the same trip id \((t_r)\) forms a set of connections representing a vehicle’s movement from the first departure station to the end arrival station. A sequence of stations includes the first and last stations along with corresponding arrival time \((t_{arr})\) and departure time \((t_{dep})\) in each station defined as a trip. A trip that covers \(n\) stations has \((n-1)\) connections. In public transportation, the vehicles that go on a trip (onward trip) and return to the starting station (return trip) are considered two different trips.

**Definition 2.3.**

A trip is a sequence of connections

\[
trip = (c_1, c_2, c_3, \ldots c_{n-1})
\]  \hspace{1cm} (4)

where
- \(c_i = (s_i, s_{i+1}, t_{dep(i)}, t_{arr(i+1)}, t_r)\)
- \(c_{i_{arr}} = c_{i_{dep}}\) and \(c_{i_{arr}} < c_{i_{dep}}\) for every \(i\)

\hspace{1cm} (5)

The footpath is considered transfer time, which is the time needed to walk between a nearby station from one trip to another trip (trip transfer). Each footpath consists of two stations with an associate the duration of walking. In a simple CSA model, transfer time has been assumed to be zero. Meanwhile, the fact is the transfer time is non-zero. The main characteristic of transfer time is often dependent on a connection, the central hub.
or station where the transfer between vehicles occurs, and the needed time. The CSA algorithm is described in Fig. 1.

```
1: Input: Connection Stream (CS), Station Stream (SS), query (sdep, sarr, tstart)
2: Output: Itinerary (I)
3: Initialize: IC, d
4: for all si ∈ SS do
5:   IC[si] ← null
6:   d[si] ← ∞
7: end for
8: d(sdep) ← tstart
9: Main Loop
10: for all c ∈ C do
11:   if tdep[c] ≥ d[sdep[c]] and
12:      tarr[c] < d[sarr[c]] then
13:      d[sarr[c]] ← tarr[c]
14:      IC[sarr[c]] ← c
15: end if
16: end for
17: I ← Itinerary [IC]
18: return I
```

Fig. 1. Algorithm CSA [22]

2.2. Methodology

The initial stage of this research is defining the problem. The problem statement is the unavailability of integrated public transportation information to support FITs to reach tourism destination. This problem leads to the objective and scope of this research: as building a seamless way journey planner that could attach to the existing official website with the scopes are focus on international tourists who enter Taiwan from Taiwan International Airport (this considered as a gateway or main hub) and the top three tourism destination, namely Jiufen (九份), Sun Moon Lake (日月潭), and Kenting National Park (墾丁國家公園).

The third stage is collecting the data. The data needed in this research is the public transport timetable data, which have a trip to the selected tourism destination. The public transport timetable data involve three transportation modes: Taiwan High-Speed Rail (THSR), Taiwan Railways Administration, and Intercity Bus. Like nature, these lead to data integration among three different transportation modes called intermodal transportation.

These data are as an input to the Algorithm Implementation Simulation. In this stage, the timetable needs to formalize and encode what transportation mode exists, the station where they depart and arrive when they depart and arrive, and how the traveler transfer between transportation modes. The timetable formalization process consists of extracting and transforming data. Once the algorithm simulation worked, the next step is the journal planner design and implementation. This stage consists of developing and testing the system’s software. As we addressed this system to be able to complement existing tourism website, the system implemented in “PHP”, “HTML”, “CSS”, “JavaScript” as a programming language and MySQL as database service to store the transportation timetable data. The research workflow of tourism intermodal transportation journey planner is shown in Fig. 2.
The tourism intermodal transportation journey planner architecture in Fig. 3 shows how the journey planner implementation process from collecting the data through the journey planner works. The users or tourists used the web application as the user interface and inputted the origin station, destination station, and departure time. In the backend, the CSA proceeds the input and gives the output as the fastest journey (the earliest arrival time route) back to the web application.

3. RESULTS AND DISCUSSION

3.1 Timetable Formalization

The journey planner works with a data structure previously built from each timetable transportation data to the CSA timetable. The timetable transportation data is being processed and transformed on a desktop computer. The processes consist of two phases, namely extracting data and transforming data.

b. Transforming Data

In the second phase, the collected data transformed into a spreadsheet, and adding the transfer path's connection with the total for real connections is 1711. After completing the spreadsheet data and importing it to the database form, the database functionality was added to the web page.

3.2 Algorithm Implementation

Before going through the system implementation, the basic form algorithm implementation is used as simulation. Start with the formalization of the table and assumption that used. The table format that we need is the connection stream and station stream. In a simple connection model, the transfer time has been assumed to be zero. The most straightforward problem that we consider to solve is the earliest arrival problem given the input of timetable, origin station \( S_{ori} \), destination station \( S_{des} \), and starting time at origin station \( t_{start} \) to find the fastest way to arrive at the destination.

![Transportation network for initial result](image)

Fig. 5. Transportation network for initial result

The transportation network describes in Fig. 5. There are four routes and eight trips in the transportation network with the different public transportation services represented by the network's different colors. The first route is Taoyuan - Taipei, the second route is Taipei - Ruifang, the third route is Ruifang - Jufen, and the last route is Taipei – Puli – Sun Moon Lake. By looking at the examples network, Taipei Station could be categorized as a transit station. It handles a substantial amount of traffic and connects the elements of the same of different transportation networks in scale. An example of the problem is the fastest way to Sun Moon Lake from Taoyuan, leaving at 06.00. Building connections are the fundamental unit building block to calculate the earliest arrival time given that timetable. The amount of connection of

---

http://dx.doi.org/10.30656/jsmi.v4i2.2772
each trip could be calculated by subtracting 1 from the amount of station, e.g., trip 07 and trip 08 have 3 stations, then the number of connections is 3 – 1 = 2. For trip 07, the connection would be like below:

\[ c_1 = (\text{taipei station}_{\text{dep}}, \text{puli station}_{\text{arr}}, 0630_{\text{dep}}, 0830_{\text{arr}}, \text{tr07}) \]

\[ c_2 = (\text{puli station}_{\text{dep}}, \text{sunmoon lake station}_{\text{arr}}, 0830_{\text{dep}}, 0900_{\text{arr}}, \text{tr07}) \]

The connection (C) stream sorted by departure time shows in Fig. 6. The basic idea how the CSA works is by keeping the best arrival time \( d \) (earliest arrival time) and arrival connection IC (InConnection) for every station (s) to get the itinerary leaving from origin station (\( s_{\text{ori}} \)) to destination station (\( s_{\text{des}} \)) at starting time (\( t_{\text{start}} \)).

![Fig. 6. Sorted connection stream](image)

The connection scan iterates all the connections to see whether it can improve any connection. The earliest arrival time from the origin station to any other station is computed by exploring all the connections. The backtracking search is utilized from the destination station and rally on the IC (InConnection) until it reaches the origin station. This is shown in Fig. 7.

![Fig. 7. The itinerary initial result](image)

After the simulation, the system implementation process is integrated with the database and web application as the user interface. A journey planner's main function is to plan one route from a starting point to a destination. Before the route planning, the user can set parameters that are considered by the route planner. The parameters are the departure station, arrival station, and leaving time. For example, the user will go from Airport Terminal 2 Station (A13) as departure station to Kenting (Xiaowan) as arrival station and at 9.14 am as leaving time (Fig. 8).

![Fig. 8. Journey planner form](image)

The journey planner lists the complete journey planner for each station, time, and transfer path between transportation modes. The transfer path helps the user to understand how many transportation modes they will use for the chosen journey. The transfer path is calculated using the best knowledge of the footpath. It shows in Fig. 9.

![Fig. 9. The tourism intermodal transportation journey planner result](image)

4. CONCLUSION

This research is providing tourism intermodal transportation journey planner in a study case of Taiwan tourism. The journey planner is successfully built in PHP language embedded in the existing Taiwan Tourism Website. Our efforts will help the international tourists get full information about the trip journey as an idea of a seamless connection.

This research does not limit to selected transportation data. It is the beginning in demonstrated the whole process of implementing the
journey planner. The scalability, manpower, and government or service companies can gather precise and updated data in the future. For example, the data could be dynamic due to some tradeoffs, such as the unique transportation table due to some special events (Chinese New Year, Summer Holiday, or Winter Holiday).

We expect to add some features to increase user value in practice later. For example, we may cooperate with transportation service operators or government transportation institutions to increase the data variety. We may add a function of minimum transfer pattern [22], add a function of transfer pattern combined footpath, cab, and bike as a transportation method in the country [23]. We may add a multiple-day trip [24] and add some art features to get a fascinating user interface design.

REFERENCES


