都市の持続可能性に向けた旅行行動と知的移動データ統合に関する包括的研究

Muroran IT

Muroran Institute of Technology

DISSERTATION:

A COMPREHENSIVE STUDY OF TRAVEL BEHAVIOR AND INTELLIGENT MOBILITY DATA INTEGRATION TOWARDS URBAN SUSTAINABILITY

Tosporn Arreeras

A COMPREHENSIVE STUDY OF TRAVEL BEHAVIOR AND INTELLIGENT MOBILITY DATA INTEGRATION TOWARDS URBAN SUSTAINABILITY

by

Tosporn Arreeras

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Engineering

Examination Committee:	Prof. Nakatsugawa Makoto Assoc. Prof. Watanabe Shinya Assoc. Prof. Mikiharu Arimura (Chairman)
Nationality:	Thai
Previous Degree:	Bachelor of Transportation Engineering Suranaree University of Technology, Thailand Master of Civil Engineering King Mongkut's University of Technology North Bangkok, Thailand
Scholarship Donor:	Monbukagakusho Honors Scholarship for Privately-Financed International Students, Japan Student Services Organization (JASSO) Scholarship Muroran Institute of Technology Scholarship Mae Fah Luang University Graduate Scholarship

Mororan Institute of Technology
Division of Sustainable and Environmental Engineering
Japan

March 2020

Acknowledgements

At first, I would like to express my sincere gratitude and respect to my academic supervisor Dr. Mikiharu Arimura, Associate Professor of the Department of Engineering. He provided continuous support, patience, motivation, passion, enduring supervision, and immense knowledge that helped me successfully carried out this assignment. Without his kind direction and proper guidance, this study would have been a little success. In every phase of the project, his supervision and guidance shaped this report to completed correctly. I am also grateful to Dr. Takumi Asada, Assistant Professor of Department of Engineering, co-supervisor, who gives me academic support, friendship, and encourages me to fight against a hard time during living in Japan. I would like to thank you and appreciations for Dr. Nakatsugawa Makoto, Professor and Dr. Watanabe Shinya, Associate Professor for their supporting and attending as a member of examination committee.

I would like to express my special thanks of appreciation to my Civil and Transport engineering teachers for their excellent delivery and kindness supporting Dr. Suksun Horpibulsuk, Professor and Dr. Sittichai Seangatith, Associate Professor of School of Civil Engineering, Suranree University of Technology. Also, appreciation to Dr. Kittichai Tanasubsin, Associate Professor, Dr. Supornchai Utainarumol, Associate Professor, and Dr. Terdsak Rongviriyapanich, Associate Professor of Faculty of Civil Engineering of King Mongkut's University of Technology North Bangkok for their support and encourage me for academic and life.

Sincere gratitude is extended to the Japan Student Services Organization (JASSO) and Mae Fah Luang University for providing the scholarship. Thankful to the Muroran Institute of Technology for giving me the excellent opportunity to study and providing a fulfilling experience. I would like to express my appreciation to all colleagues, especially Sustainable Urban and Transportation Laboratory members at the Muroran Institute of Technology for their valuable help, support, and friendship. To the Center of International Relations staffs with all support and excellent assistance since the day one, my heartfelt gratitude and appreciation to all of you. My gratitude also to friends and families for their friendship and support in Muroran and the rest of Hokkaido, that I have been visited.

Finally, I would like to acknowledge with gratitude the support and love of my family – my parents, Group capt. Sommai and Wilaiwan Arreeras; my brother, Saharat Arreeras; my sister, Sutheerawan Arreeras. Special thanks to Suppharutchaya Chongutsah for beside me during a hard time. They all kept me rolling, and this book would not have been possible without them.

"In the memory of my mother"

Topic:	A Comprehensive Study of Travel Behavior and Intelligent Mobility Data Integration Towards Urban Sustainability
Author:	Tosporn Arreeras
Degree:	Doctor of Engineering
Course:	Course of Advanced Sustainable and Environmental Engineering
Advisor:	Associate Professor Mikiharu Arimura

Abstract

According to sustainability, the trend is spreading out around the world for past decades. There are many area subjects involved, such as city planning, transportation planning, and so on, because people realized human activities harmful to the environment by consuming natural resources with less efficiency process or damage environment by social and economic movements. Currently, emerging technologies considered for the proactive procedure in extensive study areas regarding new technology application and knowledge based. In term of transport and urban study, including tourism concerns, we used intelligent data from deferent sources to be demonstrating the possible solutions which involve sustainable urban development concept. In this study, as a method of utilizing data that contributes to the sustainability of cities and regions, consideration of attractive destination management for tourists by using wireless probe data, and the weather impact on travel time reliability observation by using electronic toll collection probe data, it represented as combination experiments throughout comprehensive study. This dissertation addressed three contribution studies to the composed acknowledgment of urban mobility, and it obtained the intelligent data and specific method of research-based. It consists of; 1) an association rule mining-based exploration of travel patterns in wide tourism areas using a Wi-Fi package sensing survey, 2) Attractive destinations mining towards massive tourism area sustainable development on Wi-Fi tracking data, and 3) Assessment of the impact of snowfall on travel reliability considering different road types using ETC2.0 probe data. Hence, a stack of varying viewpoints researches provided a comprehensive review and suggestion throughout significant results. The contribution of this dissertation could be an advantage substance for strategy and policies planner to recognize alternative solutions leading to a better society.

Keywords: Association rule mining, Urban mobility, Sustainability, Tourism, Urban development, Intelligent transport system.

題目:	都市の持続可能性に向けた旅行行動と知的移動データ統合に関する包括的研究
氏 名:	トスポーン　アレーラス
学位:	博士（工学）
コース:	先端環境創生工学コース
主任指導教員氏名:	有村　幹治 准教授

論文内容の要旨

過去数十年にわたり世界中で都市の持続可能性がトレンドとなり研究対象となっている．人々は，非効率な天然資源の消費や社会経済活動による環境破壊など，地球環境に有害な活動を行い，これには都市計画や交通計画を始め，多くの分野が密接に関係している．現在では，これらを解決する新技術の開発や応用が広範囲な研究分野で日々取り組まれている．本研究では観光に関する問題を，交通と都市の研究の観点からさまざまなビックデータを使用し，持続可能な都市開発を目標とした具体的な解決策を示した．本研究では都市や地域の持続可能性に資するデータの活用方法として，Wi-Fiパケットセンサーを使用した旅行者にとって魅力的な観光目的地マネジメントに関する研究，およびETCプローブデータを使用した旅行時間の信頼性の観測における天候の影響に関する分析を組み合わせて示した．本論文では，都市の移動性の認知に対して以下に示す3つの研究から，特徴的な結果と有効な分析手法を確立した．1）Wi-Fiパッケージセンシング調査を使用した，広域観光エリアでの周遊パターンのマイニングベースの関連法則の調査，2）Wi-Fi追跡データでの大規模な観光地の持続可能な開発に向けた魅力的な目的地の抽出，3）ETC2.0プローブデータを使用して，様々な道路タイプを考慮した旅行の信頼性に対する降雪の影響の評価．以上の研究から，複数視点の考察を積み重ね，包括的な評価と提案を行い，いくつかの重要な結果が得られた．この論文の貢献は，より良い社会への問題解決への糸口となり，今後の政策立案者にとって有意義な内容となるだろう．

キーワード: アソシエーションルールマイニング，都市交通，持続可能性，都市開発，高度道路交通システム，観光．

Contents

Title		**Page**

List of Figures

List **Page**

List of Tables

Chapter 1

Introduction

1.1 Background

Since, the new technology is called the intelligent transport systems (ITS) rules the society and change the way we live on travel activity basis [1]. Soon, ITS systems will arrive that offer fundamental breakthroughs in safety, congestion reduction, driving comfort, and environmental friendliness, bringing them to levels far higher than those provided by current road transportation systems. ITS is ruling the new era of people mobility around the world [2]. The major cities have been facing similar problems on people day-life activities, regarding the urban mobility situations become worse. Traffic congestion is one of classical problem since vehicle developed and hit on the road multiple decades [3]. The congestion has changed our living today. Besides, the tourism industry as same as accessibility problems also influencing the way we live, aligning development and management protocol is essentials. Because of the number of travelers increasing throughout the year, the case of an inbound visitor to Japan is increasing excess 30million visitors in 2018, see in Figure 1.1. The number of inbound tourists is increasing annually due to tourism promotions and policies, including unique terrain, natural eco-tourism, and outdoor activities. Thus, the direction of tourism management must be considered carefully; impact on existing urbanization and the environment must clarify. The sustainability-based approach has been proposed on multidisciplinary interests over the years. It is one of the famous directions to establish an urban mobility strategy for a better life quality viewpoint.

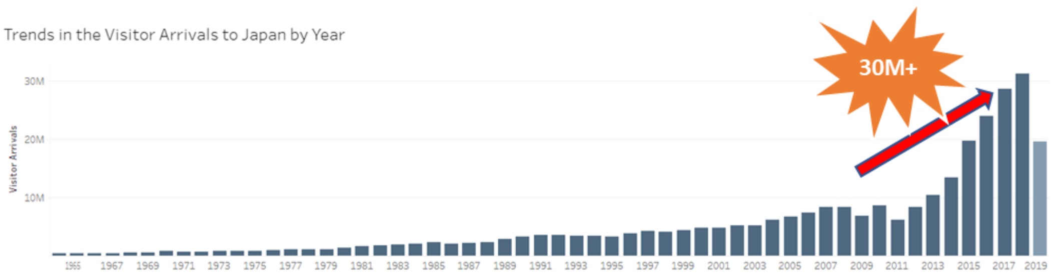


Figure 1.1 Trends in the visitor arrivals to Japan by year [4].

To explore an existing occurrence, opportunities, and guidance of research. In the starting point of the study, we used the bibliography mining method to clarify the relationship of any individual keyword, also looking for any possibility to establish research under appropriate goals and objectives. Thus, commercial software such as VOSviewer [5] is proposed to be the mining of relation and coordination of various subjects under urban sustainability provision. The software tool used to construct and visualize bibliometric networks. These networks may, for instance, include journals, researchers, or individual publications, and they can be built based on citation, bibliographic coupling, co-citation, or co-authorship relations. It also offers text mining functionality that can be used to construct and visualize co-occurrence networks of important terms extracted from a body of scientific literature. The visualization of co-word occurrence shown sustainable development, sustainable tourism, tourism, and management were synchronized areas see in Figure 1.2. Moreover, when consider on periods of trendy of subjective, the sustainable tourism and management topics seem to be a famous topic over the time see in Figure 1.3. Climate change, ecotourism, and protected areas tend to be the new position of sustainability study. The density visualization formed to simplify the most occurrence areas as shown as Figure 1.4.

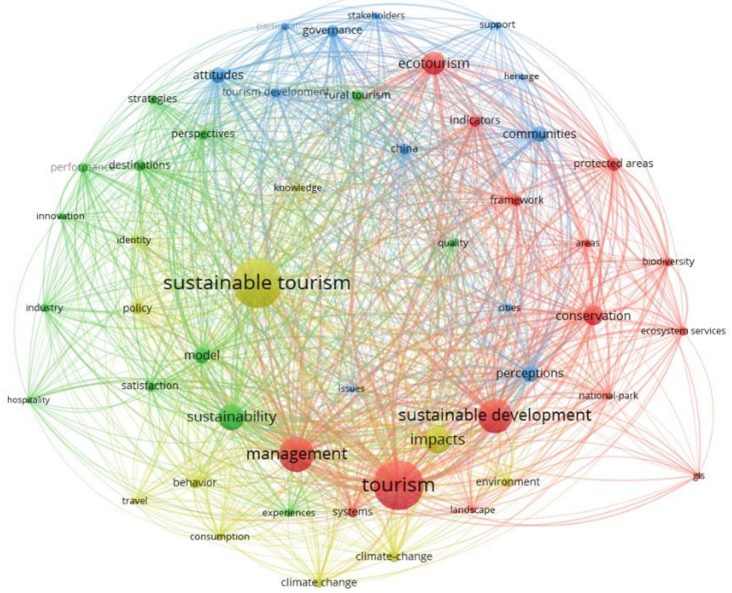


Figure 1.2 Co-word occurrence network visualization

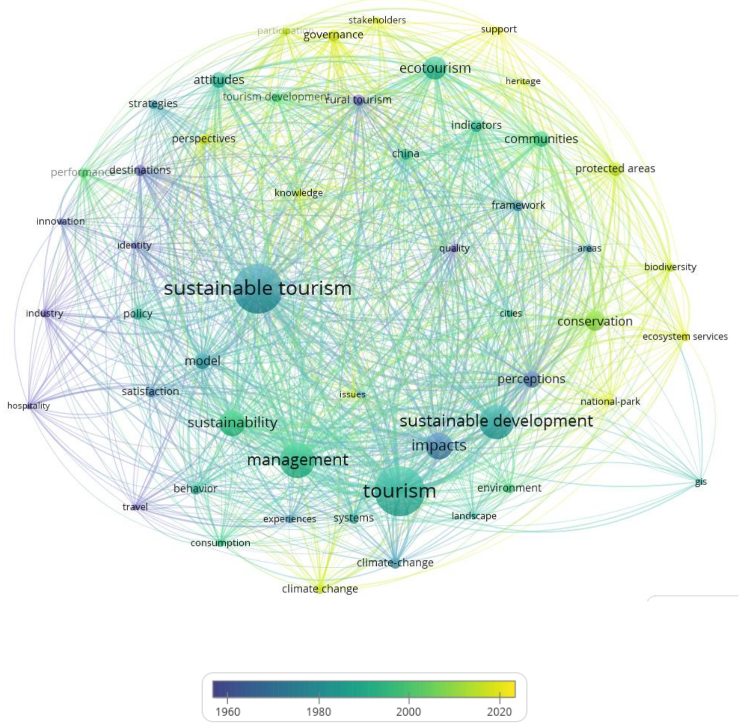


Figure 1.3 Overlay visualization consider on period over past decades.

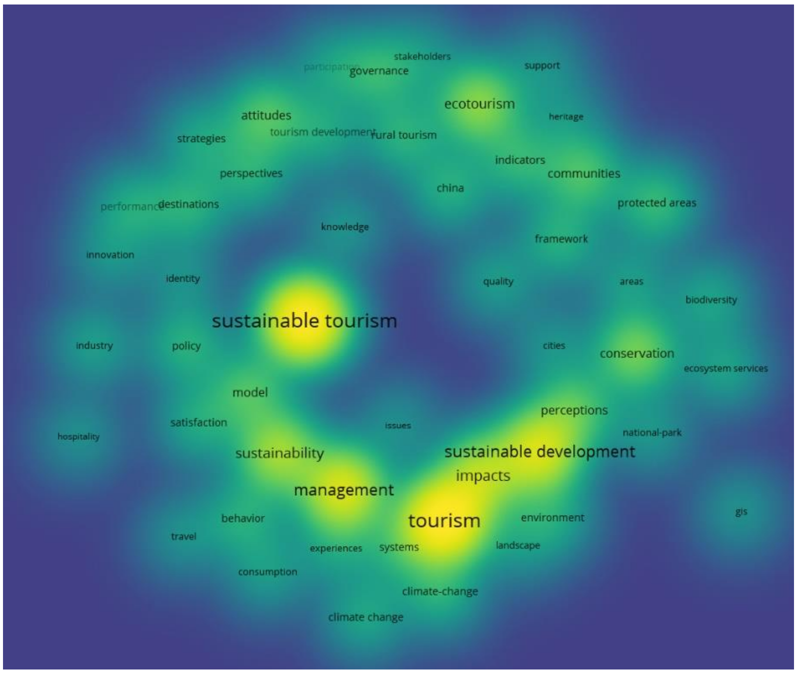


Figure 1.4 Density visualization of interested areas.

The research context is representing sustainable urban development, which coverage several ideas of interested, regrading on hand dataset and methodologies perspective. Thus, a combination of different viewpoints in this research will make a more clearly understand of urban mobility and sustainability, the research composition under sustainable urban development conceptual see Figure 1.5.

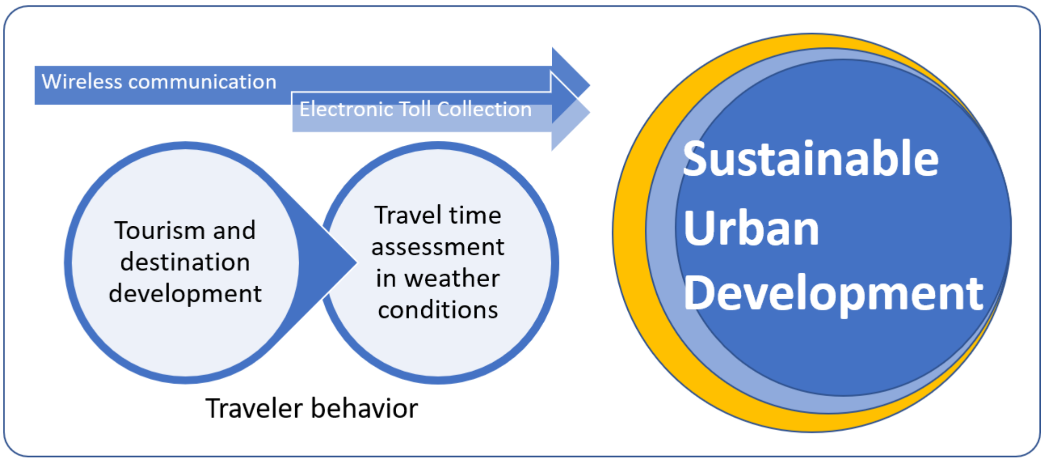


Figure 1.5 The research composition under urban sustainable development conceptual.

Accordingly, sustainable tourism consideration, each sector for tourism has its own issues in relation to sustainable tourism. There is also the question of business tourism and sustainable tourism see Figure 1.6. According to the theme of this research several issues must to be considered for sustainability approach such as transport, destination management, and visitor attractions. Transport is essential concern, because it's generating population of visitor to destination as well as pollution producing on traffic congestion, fuel used, environment interruption. Destination management including planning and development strategies, public sector subsidies and regulation of the tourism industry, financial, and tourist taxation. The last, visitor attractions are considered on environment impact of new infrastructures and operations also play attention on relation with the local community. Hence, we emphasize how the improvement of market basket analysis adopting association rules is appropriated to determine destination management problems with tourism accordingly. Moreover, traffic problems such as congestion and delay regarding weather conditions have proposed the travel time reliability considerations. It was a scenario of urban mobility, including a sustainable tourism issue representative.

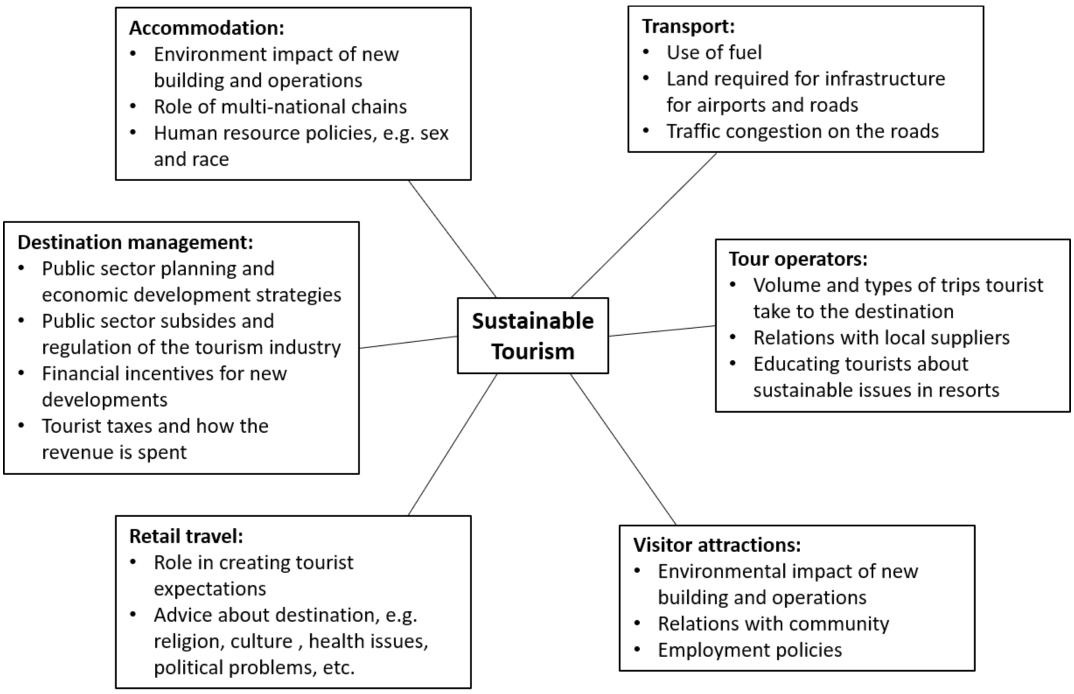


Figure 1.6 Sustainable tourism issues and the different sectors of tourism [6]

1.2 Objectives and Research Questions

The combination of studies under single similar concept of urban mobility investigations. The main research questions have proposed to be challenging the scientific achievement procedures, also magnified understanding of urban mobility problems from several point of under sustainability provision. There are represented data mining technique such as association rule analysis to determine the traveler behavior that related to attractive tourist destination. In this research, we simplified the main objective as follow;

“To amplify urban mobility development through association rules mining and wireless probe data application enhancement, towards sustainability approach.”

To achieve the objective of this research. Several questions established to full fill gaps of the study also clarify the enhancement of market basket analysis for understanding travel behavior. Moreover, technology integration and discussion are challenging for future urban mobility studies. The research questions as follow;

Question 1: How to enhance the urban mobility atmosphere following sustainability procedures?

Question 2: How association rules mining can apply to destination development significantly?

Question 3: The wireless technology advantage is practical for urban mobility investigation?

1.3 Contributions and Methodologies

The contribution of this research reviewed according to the results of the subsequent chapters. The mechanism of analysis supervision on this research is simplify represented see Figure 1.6.

- **Chapter 2:**

  An Association Rule Mining Based Exploration of Travel Patterns in Wide Tourism Areas using A Wi-Fi Package Sensing Survey.

- **Chapter 3:**

  Attractive Destinations Mining Towards Massive Tourism Area Sustainable Development on Wi-Fi Tracking Data.

- **Chapter 4:**

  Assessment of The Impact of Snowfall on Travel Time Reliability Considering Different Road Types Using ETC2.0 Probe Data.

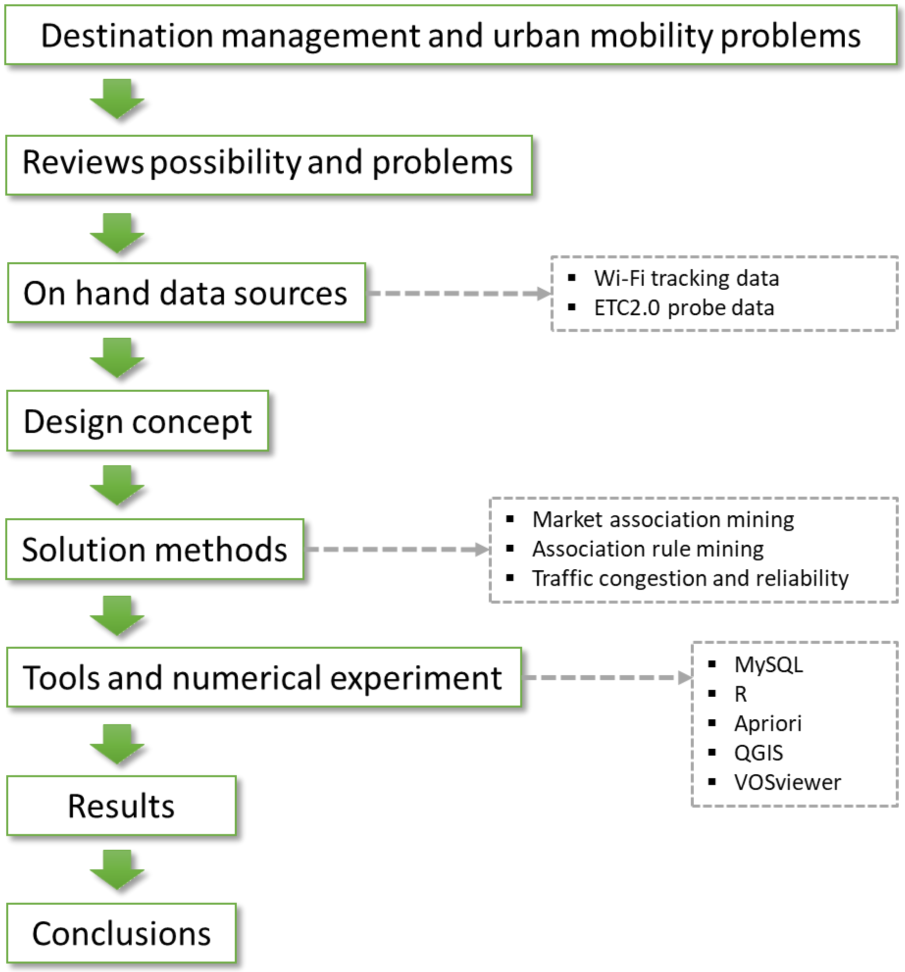


Figure 1.7 Main conceptual framework

1.4 Outline of Dissertation

This research consists of six chapters, which contain three major research as series of integration of intelligent transport systems to city mobility and accessibility assessment study. The section providing details of research introduction and recognition of study individually as follow;

Chapter 2: The chapter is proposed the market basket analysis (MBA) capability in tourism study. Since many articles over decades rare apply into the field study of urban mobility or tourism management approach. It provided the opportunities to setup this experiment involves wireless probe data from the tourism area, demonstrated in one the most attractive travel region in Japan, Hokkaido. Since, the number of inbound tourists in Hokkaido is increasing annually due to tourism promotions and policies, including unique terrain, natural eco-tourism, and outdoor activities. The attractive destinations and journeys information can assemble to predict traveler mobility. Distance and travel times affect travel behaviors in Hokkaido; also, most of the attraction locations are in the countryside and away from the city center. Hence, travel patterns are considered to maintain and improve the tourism atmosphere. In this research, travel patterns analyzed employing Wi-Fi probe data, which was collected in Asahikawa and Furano tourism areas. First, fundamental analysis was conducted at each spot. Next, the three indicators support, confidence, and lift parameters were calculated using the Apriori algorithm, which generally is used in marketing purposes. The results, the usefulness of association rule mining calculation provided rules of the transaction of migration travel characteristics — sequential travel patterns illustrated to identifying significant location toward sustainable tourism development.

Chapter 3: This chapter adopted previous chapter methodology, the association rule mining analysis to empowered attractive destination identification processes. The challenge of intelligent data integration with appropriated algorithm as Arpiori, it’s bought the new perspective of urban mobility investigations and tourism study, which advantage and practical for destination management procedures. This study used a Wi-Fi scanner device to track tourists’ traveling behavior in Hokkaido’s tourism area, which occupies a large region that features a unique natural landscape. Inbound tourists have significantly increased in recent years; thus, tourism’s sustainability is considered to be important for maintaining the tourism atmosphere in the long term. Using internet-enabled technology to conduct extensive area surveys can overcome the limitations imposed by conventional methods. This study aims to use digital footprint data to describe and understand traveler mobility in a large tourism area in Hokkaido. Association rule mining (ARM)—a machine learning methodology—was performed on a large dataset of transactions to identify the rules that link destinations visited by tourists. This process resulted in the discovery of traveling patterns that revealed the association rules between destinations, and the attractiveness of the destinations was scored on the basis of visiting frequency, with both inbound and outbound movements considered. A visualization method was used to illustrate the relationships between destinations and simplify the mathematical descriptions of traveler mobility in an attractive tourism area. Hence, mining the attractiveness of destinations in a large tourism area using an ARM method integrated with a Wi-Fi mobility tracking approach can provide accurate information that forms a basis for developing sustainable destination management and tourism policies.

Chapter 4: Presents the investigation on weather impact on travel time reliability (TTR), that focused on winter and nonwinter period of thought out the year. Since, the intelligent transport system (ITS) provides new idea for making better urban life by high-technology system for decades. The study aims to working with the latest version of electronic toll collection system, the ETC2.0 technology, its capability to captured travel behavior as a probe vehicle data. It provided an elaborated and accurate data, even greater than traditional data collections. Thus, this study considered on operation of weather conditions and road hierarchy configuration by using travel time-based performance evaluation, the travel time reliability related to weather conditions under the road hierarchy represented with various time reliability indexes. The comparison analysis between during winter and non-winter conditions discussed that a remarked travel time is more reliable during non-winter, and on the other hand, unreliable travel time significantly increases during snowfall. The experiment result also shows the amount of snowfall involved travel time reliability alteration.

The final, Chapter 5: express the summary of differentiated tasks and analyses, then provided the conclusion and recommendations, including the possibility of further work. Move over, the outline of the flow of the research, as mention above, is providing an overview and make it straightforward to understand the workflow see Figure 1.8.

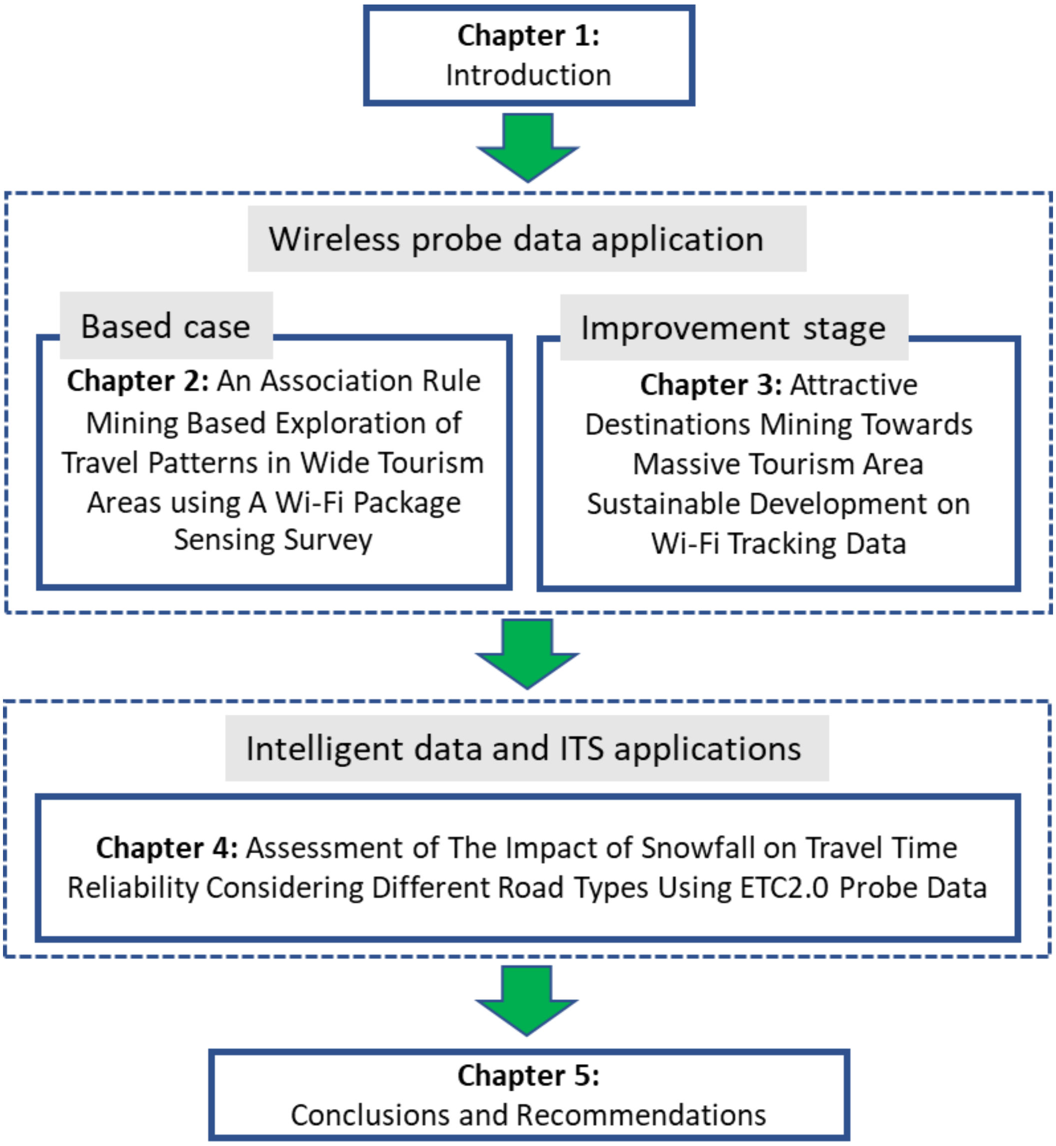


Figure 1.8 Research outline

1.5 References

[1] ITS Japan: http://www.its-jp.org/english/what_its_e/ (accessed on Nov 1, 2019).

[2] Japan's Expressway Network: https://en.driveplaza.com/expressways/maps.html (accessed on Nov 1, 2019).

[3] ITS Technology Enhancement Association: https://www.its-tea.or.jp/english/its_etc/service_etc2.html (accessed on Nov 1, 2019).

[4] Japan Tourism Statistics: https://statistics.jnto.go.jp/en/ (accessed on Nov 1, 2019).

[5] Vosviewer visualizing scientific landscapes: https://www.vosviewer.com/ (accessed on Nov 1, 2019).

[6] Swarbrooke J. Sustainable tourism management. Cabi; 2015.

Chapter 2

An Association Rule Mining-Based Exploration of Travel Patterns in Wide Tourism Areas using A Wi-Fi Package Sensing Survey

2.1 Introduction

The Hokkaido region, one of the most famous tourist destinations in Japan has had significant increases in visitor numbers in recent years [1]. There are many natural attractions in Hokkaido's 83k square kilometers landscape, which attract visitors from around the world to engage the tourism industry. This area is well known for volcanoes, natural hot springs, and ski resorts. The region is occupied of unique terrains, with widely scattered structural areas, including cities and sightseeing localities. Most of the attractive tourist areas are located in the countryside, necessitating the use of private transportation is considered. Although, visitors tend to choose the private mode such as an owned vehicle or a rental car service rather than public transportation, to reach target destinations. It is generating traffic congestions and pollutions may cause a toxic environment for local residences, particularly throughout in high tourism seasons.

Recently, there are various articles adopted trajectory investigation analyses and alternative observation techniques for tourism mobility approach. Which was usually come along with limitation problems for example survey cost, bias of results, and the penetration rate of observation devices. Also, the difficulty of applying scanner devices in large areas for long-term observation. Therefore, to understand the phenomenon of traveler trajectory in large areas, using the new observation technology such a Wi-Fi scanner instrument in field survey, it was considered as an alternative survey method. Because it can be provided significant accuracy and reliability, while compares to conventional methods.

Regarding previous tourism studies, data mining was generally employed to analyze the trajectory pattern in complex dataset; for example [2] presents a behavioral study by extracting spatiotemporal metadata from tourist's social media consuming behaviors. They applied the sequential pattern mining to extract the location of social media users have been visited, also the movement trails such time and where they are going next. It was described as useful mining technique to determine the spatiotemporal of tourist movement. Like, an associate rule mining technique adopted in tourism studies [3–5]. The articles were explored the associations of different visiting locations, also indicated the potential of using a tracking device to observe tourist's behavior. As above, data mining techniques were registered to explain the travel behaviors in tourism, which consist the classification, prediction, and association contexts. Association analysis is a data mining technique for determining the fascinating relations between the variables in the large data set, also classify information in databases accurately [6]. It was an essential and popular method in the business sectors, including medical and pharmaceutical research [7]. In transport and tourism studies, administrators can identify the relationship of visiting place to another place in order to tourism's management and planning schemes [8], including visualization diagrams used for demonstrating the visitor's travel patterns in multiple attractive locations.

This study observed visitors' digital footprint in Hokkaido's wide tourism areas, as shown in Figure 1. The observation locations were equipped multiple Wi-Fi package sensors, to catch up the radio transmission between visitor's mobile devices and providers on site. This survey method was producing necessary data such as probe requested data, which was a fundamental data of study. Then, the association rule mining method was applied to determine the relation between individual visiting locations and finding the sequential pattern in transaction data. The results would address the most attractive location, where traveler frequently visited gather on their trips. It can be explained by set of rules as the association rules.

2.2 Literature review

Travel behavior research tends to a popular topic in recent years. The survey innovation was proposing a new era of data collecting technologies instead of conventional survey methods such as a questionnaire interview on the paper-based survey. The article argued smart technology solutions able to catch up high-resolution data from observing traveler's mobility in large-scale population [9]. Challenge of data collecting by smart technology has been changed perception of survey study, since the streaming of information, communication, and technology (ICT) services are engaged the globalization. It was provided opportunities to utilize general smart communication device such as smartphones to collect fine details of traveler behaviors [10].

The challenge from the past studies, since the digital footprints [11] tracking trendy used in social mobility studies, including discussion on tracking technology and extraction processes were considered. Therefore, social media platforms were considered as an alternative probe data sources, while radio frequency detectors such as Bluetooth technology [5] plays as digital footprint extraction device from enabling wireless communications. According to various past studies, social media trace on media platform such as Twitter and Flickr was acquired for exploring traveler behavior and trajectory patterns, focused on user's smartphone in wide tourism area [9][11].

Recently, the exploration activities and mobilities of the traveler by smart sensing technologies tended to be more reliable [12], the anonymous smart sensors can capture the media access control (MAC) address from each wireless device within the scanning area. MAC address is unique to each device, and it gives direct observation on the enabled device's movement along considered routes [13]. In the business viewpoint, prediction of market shares on the new catering location by using Wi-Fi traces to detect sequential of occurred activities on campus, MAC address and username will be captured for choice of location modelling [14]. [15] introduced plug-and-play Wi-Fi sensors provision configurated for collecting the Wi-Fi probe data in order to indoor localize of human presence in specific areas. Similar to [16] discussed a key benefit of MAC address as an emerging technology; it is possible to track the spatiotemporal movement of a human in term of space utilization.

According to related articles are mainly point out the small space utilization such as a room in campus building even the campus area, there are few studies related to large area observation as shown as the above section. Thus, extensive area case studies are considering. [17] estimated the origin and destination for city transit in a wide area. Moreover, [10] investigated the paratransit passenger boarding and alighting locations in specific tourism area by using Wi-Fi scanner device based on Raspberry PI, the unit cost is inexpensive, and possible to install in several places. The Wi-Fi probe requests [18] can be utilized to estimate the waiting time of passenger at the bus terminal, it was generating from traveler's connected Wi-Fi devices such as a smartphone. From different studies, MAC address, was played as a digital footprint social movements observation. Passive Wi-Fi tracking process as a digital footprint was conducted for a single mobile phone localization detection on the massive spatiotemporal trajectory, similar to global positioning system trace for detected wireless communication devices [12].

The extraction of the extensive trajectory database, the sequential pattern mining trendy deployed to transport travel sciences [2]. Moreover, to understand the traveler behavior, the pattern mining techniques is an appropriate mechanism to discover the hidden patterns in large databases [19]. The market basket analysis with its well-known algorithm as an Apriori algorithm [20], the machine learning techniques [21] used for identifying the relation between items. It addresses the sequence of items that occur together frequently in transactions. In the other hand, it empowers marketer to classify bonds between the items that people purchase or consider. In the case of this study, the abilities of association mining method would identify the individual behavior of traveler in specific areas, regard the rare case of applying association mining technique application and the market basket concept to understand travel behavior and contribute to tourism planning and management with tourism study.

2.3 Methodology

2.3.1 Choice of Location and Sensor's Placement

This research conducted on specific target areas, where involved with traveler mobilities and activities in a large tourism areas of Hokkaido region. The Wi-Fi sensors were installed in 31 different places at the attraction traffic node including the SA / PA (service area and parking area) on the expressway, highway stations, major tourism/commercial facilities, etc. as shown in Table 2.1. In case of various major facilitates located at the single area multiple devices were installed, — the investigation period, total 33 days from June 19, 2017, to July 23, 2017. According to Furano and neighbor areas, as shown in Figure 2.1, these places attract many tourists due to their beautiful flower gardens, especially the lavender plantation field during blooming period from June and July.

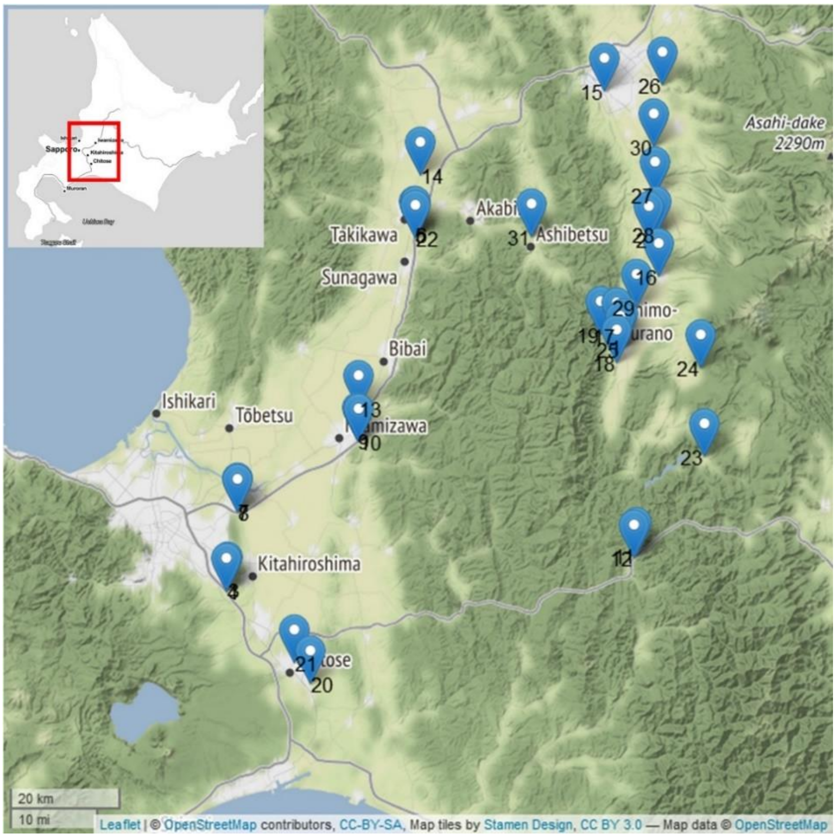


Figure 2.1 Geographic of central Hokkaido including Asahigawa and Furano tourism areas, all 31 observation spots marked in blue.

Table 2.1 The list of observation spots in tourism area.

No.	Location	No.	Location
1	NaturaxHotel	17	CampanaRokkatei
2	LavenderOwnerGarden	18	FuranoCheesePlant
3	PAWattsuUp	19	LavenderHighlandFurano
4	PAWattsuDown	20	ShinchitoseTOYOTA
5	PASunagawaUp	21	RSSalmonpark
6	PASunagawaDown	22	SAHighwayoasis
7	PANopporoUp	23	RSMinamifurano
8	PANopporoDown	24	Gorohouse
9	PAIwamizawaUp	25	FuranoMarche
10	PAIwamizawaDown	26	AsahiyamaZoo
11	PAShimukappuUp	27	RSBieinokura
12	PAShimukappuDown	28	Shikisainooka
13	RSMikasa	29	NakafuranoLavenderGarden
14	RSTakikawa	30	AsahikawaTOYOTA
15	RSAsahikawa	31	RSStarPlazaAshibetsu
16	FlowerLandKamifurano		

2.3.2 Wi-Fi Tracking Instrument

The detector consists of Raspberry Pi 2; it is an operating system hardware integrated with several components; the external USB antenna, powered by portable batteries pack or standard USB charging cable, and SD memory card, see Figure 2.2. The device is developed for detecting Wi-Fi transmissions and capable to capture probe requests data automatically, which emitted continuously from enabled Wi-Fi connection device such as a smartphone. The operational range can coverage 300 m to 400 m of radius distance under ideal conditions. Hence, the software installed to enable the device to collect probe requests then converts into a simplified format. The probe request intervals record is entirely dependent on the individual emitting source, and the communication function makes possible of transmits data to server at the 5-minute interval (some sensors equipped with Wi-Fi router that capable real-time data transmission).

Figure 2.2 The Wi-Fi packet sensing device.

2.3.3 Data Processing and Reduction

The probe request cleansing process carried out by the following procedure.

Data integration of multi-location sensor: Although the reception range of the Wi-Fi sensor packet can coverage 300-400 m of radius distance, reception of probe request from user can be varied due to critical environment with obstacles such as buildings, crowds, weather, etc. For that reason, multiple sensors installation was required in a large area and critical environment. In the analysis, the received data at the plurality of locations were combined as the data of the spot.

Data extraction on arrival and departure of spots: In the original data, received records continuously at intervals of several seconds, but except for recording at the time of initial reception (arrival) and final reception (departure) in the spot, they were unnecessary and excluded.

Data exclusion with ten days or more excursion: Arrival and departure data for each of the above spots arranged in chronological order of each MAC address, and the round-trip time of the user was obtained from the difference between time of arrival at first spot and departure at the last spot. Users with long period movement more than ten days in the observation area can be considered as local resident or long-term stay were excluded.

2.3.4 Association Analysis

The association rule mining (ARM) is a rule-based machine learning method, and it is used for determining the satisfaction patterns of relations between variables in the extensive database. Association rule algorithms that well known and currently used is the Apriori algorithm [21]. Also, the association rules are an essential class of regularities in data, and it is a fundamental data mining process to finding all co-occurrence relationships among data items [22]. The association rule mining, also known as the classic application in market basket analysis, it widely applied for analysis as data mining method, which designed to identify group of variables that are highly correlate to each other [23]. However, to obtain the best rule of ARM, the principle of this method consists of three measure elements: the support, confidence, and lift of the rule. There are often used for the interestingness of association patterns evaluation [24] in the element's framework as follow.

The support of the rule is indicating how frequently of the items collection perform together, while N is the number of total transactions, which computed by Equation 2.1. Support is a useful measure; the high value of support represents the frequency of repeated, which means the itemset was selected more often. Thus, the relevant aspect of itemset should have high support value nor low value because obtaining low value may not profitable. The rule may occur due to chance; it also considers as an inappropriate rule [25].

$$Support\ (X \rightarrow Y) = \frac{Count\ (X \rightarrow Y)}{N} \tag{2.1}$$

The confidence of the rule is representing the indication accuracy or conditional probability of performing itemset. Its value respect to set of transactions, that contains X, which also contains Y. The value of confidence is consisting of probability between 0 and 1; thus, the association with the highest confidence value would consider as it is reliable. It computed as Equation 2.2. However, the confidence purely may not be enough to explain the results, therefore, the Lift value was applied to address the problem [26].

$$Confidence\ (X \rightarrow Y) = \frac{Support(X \rightarrow Y)}{Support(X)} \tag{2.2}$$

The lift of the rule or lift value described as a general measurement of the association between the itemset, it was represented by the confidence of the rule divided by the support. Besides, if the value equals to 1 specify zero correlation, greater than 1 specifies positive correlation, and less than 1 specifies negative correlation. Addition, the lift ratio was explained as a general measure of the association between transaction datasets; thus, higher lift ratios indicate stronger association. Therefore, the satisfaction value should be more significant than 1, and it reflected the true relation of itemset, which found gather more often than by chance [20]. It computed as Equation 2.3.

$$Lift\ (X \rightarrow Y) = \frac{Confidence\ (X \rightarrow Y)}{Support\ (Y)} \tag{2.3}$$

There are various appropriated methods that used for identifying the relations between variable as the sequential pattern mining is popularly used for extracting travel pattern [2], and the series pattern mining also considered as correlation rule analysis used in big data analysis [27]. ARM method used for exploring the association of the order of visiting spots and the representative of patrol patterns in the wide tourism area, from the wireless transmission database. It was conducted “if-then” statements for essential variable values [28]. Thus, this advantage would describe in the pattern mining utilization, also it possible to visualize the combination of location order and travel patterns.

2.4 Results

2.4.1 Implement Association Rule Mining

To performing the ARM analysis using R version 1.1.463, dataset preparations were required. The pattern of the transaction from probe requests dataset, only continuous transaction over time can be considered as transaction data. Since the mac address has recognized in duplicate number as well as date and time, the location at the last column identified as difference places, these characteristics of dataset selected. Due to mac number which identifies the person was visited the place and never been changed, then if sensors captured those transmission same mac number in different locations, it's can be considered as the movement of a specific traveler from one location to another. Table 2.2 shows an example of transaction records and appropriate continuous pattern of ARM.

Table 2.2 Example of probe requests data that use in ARM analysis.

Mac number	Date	Time	Site
00016345a532cb2286b2cf328a228e4c	626	11:20:03	RSMikasa
0001717a4ed86d994bcfec6c03c71e94	723	13:47:47	Shikisainooka
0001717a4ed86d994bcfec6c03c71e94	723	14:56:37	AsahiyamaZoo
000185331e06a73de2faa22a949d8e24	722	11:20:43	NakafuranoLavenderGarden
000185331e06a73de2faa22a949d8e24	722	11:20:57	NakafuranoLavenderGarden
00009872164fb495849d2de5f8561b14	712	9:47:03	PAIwamizawaUp
00009872164fb495849d2de5f8561b14	712	15:17:24	Shikisainooka
00009872164fb495849d2de5f8561b14	712	17:09:31	FuranoMarche
00009872164fb495849d2de5f8561b14	712	19:08:36	PAIwamizawaDown

After preprocessing, the dataset includes 1,048,575 records and four attributes, then convert data frame into transaction data. All sites with the same mac number and day of visiting was stored in one row as the single format. Additional, optional attributes have terminated; the only set of site location remains in transaction dataset called the basket format. The final transaction dataset had been applied with the Apriori algorithm to conduct the item matrix, only 440,061 transactions and 31 items of location remains. The transaction length distribution reported the travelers' behavior of visiting places and traveling pattern on their trip; about 95% of travelers visited only one place on their trip see Table 2.3. The most frequent locations by travelers were AsahiyamaZoo, FuranoMarche, Shikisainooka, SAHighwayoasis, and FuranoCheesePlant, respectively, as shown in Figure 2.3.

Table 2.3 Transaction length distribution size.

Rule length	1	2	3	4	5
Transaction	417,695	22,350	11	4	1

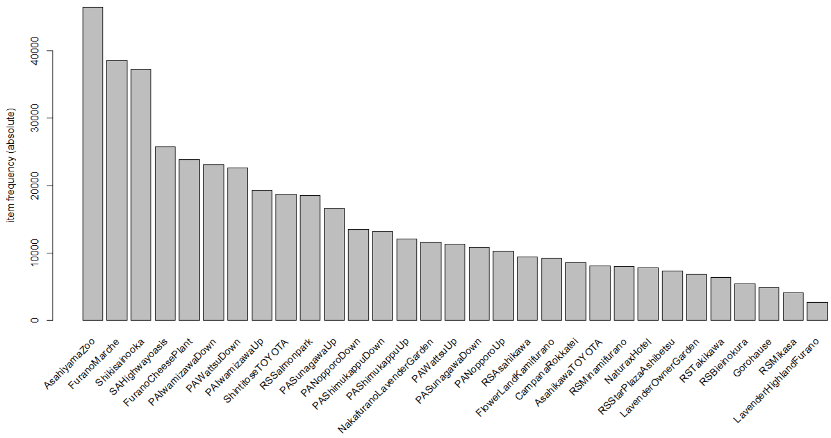


Figure 2.3 The absolute transactions by visiting location in wide tourism areas.

2.4.2 Enhanced Apriori Algorithm

The enhanced Apriori algorithm-based ARM, established minimum support and confidence required. Identify the redundancy rules process was a necessary process to sort the duplicate rules, then removed redundant rules procedure applied. Based on the calculated relationship of the three indicators, the tendency of the excursion pattern in the target area period visualized. In association rules, for extracting the meaningful transactions, thresholds were set in the degree of support, confidence, and lift. Since the number of transactions in this research was enormous, the degree of support tends to be small. Therefore, the ARM process was set a threshold of support of 0.001 and confidence of 0.01 to return all the rules, that support of at least 0.1% and confidence of at least 1%. Regarding the enhanced process, the ARM where the left-hand side (LHS) was the origin of the transaction and right-hand side (RHS) was the final observation area, after removed redundant rules there were only 8 rules left out of 16 rules with only rule length distribution of two, it was sorted by decreasing of confidence see in Table 2.4.

The rules performed a pattern of visiting places, and the highest confidence reported 6% of the traveler who visited the parking area “PAIwamizawaUp” also visited “PAIwamizawaDown”, about 3% of the traveler who visited “PAIwamizawaUp” also visited “Shikaisainooka” where was free of entry charge to flower garden in Furano. From the relationship between the above three indicators, it could grasp the whole picture of the traveling pattern in target areas. Next, to extract large transactions of each indicator, consider typical trip patterns of large sightseeing area and measures to increase the number of visitors. For the sake of clarity, in the degree of support, confidence, and lift value, the transactions were arranged in descending order of value, and the rules were extracted and visualized as network diagrams.

Table 2.4 Association rules of visiting places.

Rule	LHS	RHS	Support	Confidence	Lift
1	PAIwamizawaUp	PAIwamizawaDown	0.00252	0.05747	1.09732
2	PASunagawaUp	PAIwamizawaDown	0.00146	0.03853	0.73572
3	Shikisainooka	FlowerLandKamifurano	0.00111	0.01314	0.62519
4	Shikisainooka	PAIwamizawaDown	0.00187	0.02214	0.42266
5	PAIwamizawaUp	Shikisainooka	0.00116	0.02640	0.31213
6	FuranoCheesePlant	FuranoMarche	0.00120	0.02219	0.25321
7	AsahiyamaZoo	PAIwamizawaDown	0.00119	0.01131	0.21599
8	SAHighwayoasis	AsahiyamaZoo	0.00110	0.01887	0.17890

2.4.3 ARM Analysis

The most attractive association rules among locations are shown in previous section. If attention the tourism places instead of parking areas, Shikisainooka where was the frequent chosen by visitors, likewise the AsahiyamaZoo seem to be the most famous for visiting during the summer season. There was a high possibility of the traveler who visited Shikisainooka then visited FlowerLandKamifurano regarding to the high lift value of 0.63 compared to other rules. Moreover, the parking area visited also had high lift value as PAIwamizawaUP then visited PAIwamizawaDown had the most reliable lift value of 1.10, it shown the highest possible of traveler to consider visiting more frequently than other places.

Besides, some rule patterns made parking area rise at the final place (RHS). These were visiting patterns from Sapporo area to the Furano area for sightseeing. These were representatives and typical patterns of target users, but the confidence level and lift value were low. Therefore, patterning of the final place has not fixed, and since it considered that many travelers had already visited in the current situation, it may not be considered as the pattern for promotion in the future. Also, regarding PAIwamizawa descent as a final place, it is seemed to be a typical pattern of stopover at the time of returning to Sapporo after sightseeing in the Furano area. People who interested in sightseeing stop by this PA, it was considered as an effective tourism measures for making traveler to visit by providing tourist information of both locations and surrounding areas. To figure out the potential location in the future, the visualization techniques were powerful tools for clarification.

2.4.4 The rules visualization

Matrix-based visualization technique see in Figure 2.4, is a useful explanation method; it organizes the antecedent and consequent transactions on the x and y-axes, respectively. The selected measure, it displays at the intersection of the antecedent and consequent of given rules. There some empty cells regard many potentials of association rules will not meet the required minimum thresholds on support and confidence. The matrix-based conducted by transactions in antecedent (LHS) 1= PASunagawaUp, 2= PAIwamizawaUp, 3= Shikisainooka, 4= FuranoCheesePlant, 5= AsahiyamaZoo, and 6= SAHighwayoasis while transactions in consequent (RHS) consist by 1= AsahiyamaZoo, 2= FuranoMarche, 3= Shikisainooka, 4= PAIwamizawaDown, and 5= FlowerLandKamifurano. Above figure shown the matrix with all eight rules, antecedent considered on PAIwamizawaUp and Shikisainooka that intercept the consequent layer of PAIwamizawaDown. In the other hand, the matrix plot was predicting the consequent of PAIwamizawaDown which cross along the antecedent was the most attractive location among others. This location is a parking area on the expressway network, and it may be considered for promoting and upgrading on-site facilities to be able to support the increase of visitor in the future, as well as, PAIwamizawaUp and Shikisainooka.

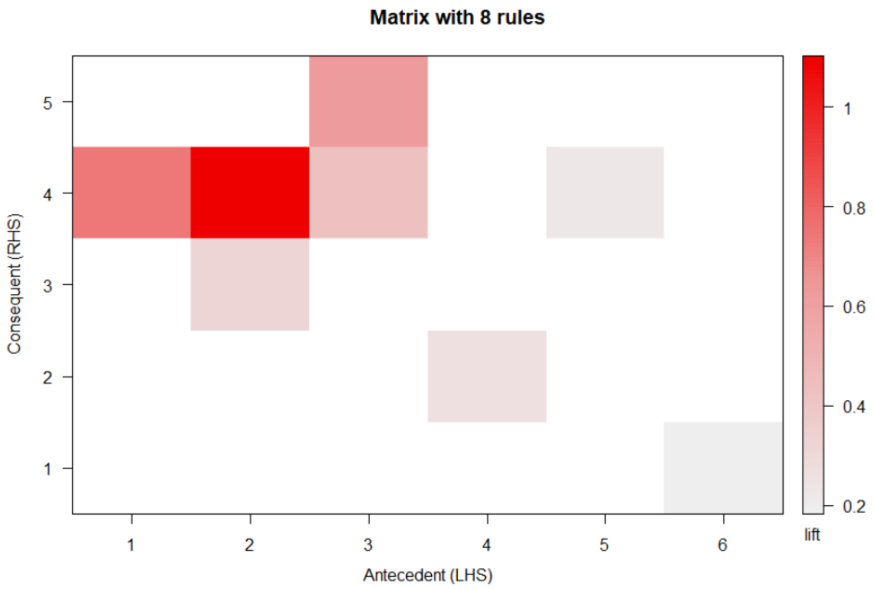


Figure 2.4 Matrix plot represent the associate rules

Likewise, Figure 2.5 visualization the association rules pattern using the graph-based technique. The rules arrows were pointed from the items to rule, vertices designated the LHS items also an arrow from a rule to an item registered the RHS. The plot can be seen at least four rules approach PAIwamizawaDown spot; this is an attractive location with contribution from others.

Moreover, there is the location with two ways contribution; PAIwamizawaUp and Shikisainooka consistent the matrix-based plot that point out as the potential location, it must be concerned for supporting the demand of travel shortly. The outstanding rule was represented traveler who visited FuranoCheesePlant then headed to FuranoMarche. Since there are many parking areas as the destination, it estimated that the increase of tourist number in the Furano area has a significant influence on the traffic condition on the expressway. Thus, it is necessary to investigate appropriate admission capacity in the parking area under the anticipation of the future increase in tourists in the Furano area.

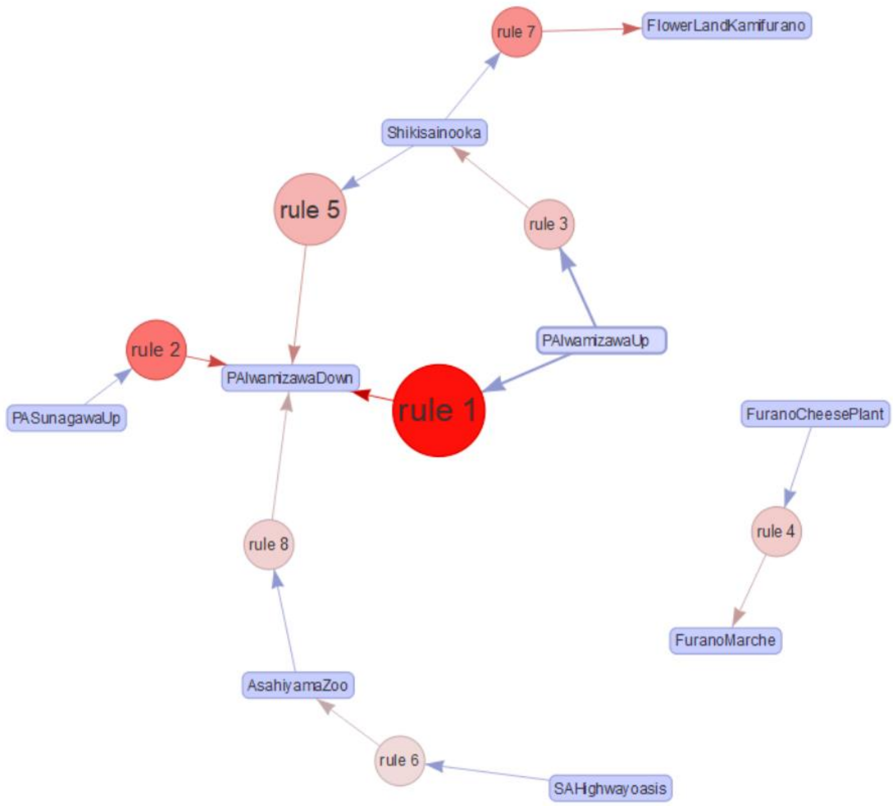


Figure 2.5 Graph plot represent association rules pattern.

2.5 Conclusions

In this research, the association rule mining applied to figure the potential location for tourism development in the wide tourism area of Hokkaido; including Asahikawa and Furano. The trajectory data in the extensive area as a probe request dataset, generated by the Wi-Fi packet sensing devices. It provided a big data database of traveler characteristics in target areas. The results, the ARM analysis could grasp the whole image of the traveling pattern in the target areas, associated with the relationship of the three indices of support, confidence, and lift. Meanwhile, the extraction of transaction dataset facilitates to capture the behavior of traveler in each location. The visualization methods of matrix rules and graph as network diagrams conducted, to represent the transactions and association rules. There was recognizing the relationship between the visiting place and final the destination, also simplify the travel pattern and useful for tourism management practically.

The results show that “PAIwamizawaDown” the expressway facility, represented as the center point destination, and unique travel patterns in this area detected. The travelers tend to visit several recreation places on their trip, meanwhile, the expressway used for reaching the targeted destinations. There is indicating the association of an individual destination with another significantly, also reflected travel behaviors in this area, factors such as weekend traveling, and seasonality tourism affects the destination selection. According to methodology, association rules mining performed under marketing concepts, its capability providing an opportunity and challenge to determine the tourism and destination problems. The integration of tracking technology in survey procedure was providing high-resolution dataset reserve for analyses. These finding could be considered as essential information for the planner, to identify the significant location for the further development plan, and decision making on promotion policies, including tourism planning and sustainable development.

2.6 References

[1] JNTO-Japan National Tourism Organisation Japan Tourism Statistics Available online: https://statistics.jnto.go.jp/en/graph/#category--6 (accessed on Jan 22, 2019).

[2] Bermingham, L.; Lee, I. Spatio-temporal sequential pattern mining for tourism sciences. Procedia Comput. Sci. 2014, 29, 379–389.

[3] EMEL, G.G.; TAŞKIN, Ç.; AKAT, Ö. Profiling a Domestic Tourism Market by Means of Association Rule Mining. Anatolia 2007, 18, 334–342.

[4] Law, R.; Rong, J.; Vu, H.Q.; Li, G.; Lee, H.A. Identifying changes and trends in Hong Kong outbound tourism. Tour. Manag. 2011, 32, 1106–1114.

[5] Versichele, M.; de Groote, L.; Claeys Bouuaert, M.; Neutens, T.; Moerman, I.; Van de Weghe, N. Pattern mining in tourist attraction visits through association rule learning on Bluetooth tracking data: A case study of Ghent, Belgium. Tour. Manag. 2014, 44, 67–81.

[6] Du, Y.; Zhao, M.; Fan, G. Research on Application of Improved Association Rules Algorithm in Intelligent QA System. In Proceedings of the 2008 Second International Conference on Genetic and Evolutionary Computing; IEEE, 2008; pp. 414–417.

[7] Lee, D.G.; Ryu, K.S.; Bashir, M.; Bae, J.-W.; Ryu, K.H. Discovering Medical Knowledge using Association Rule Mining in Young Adults with Acute Myocardial Infarction. J. Med. Syst. 2013, 37, 9896.

[8] Ritchie, J.R.B.; Crouch, G.I. The competitive destination: a sustainable tourism perspective; CABI: Oxfordshire, 2003;

[9] Hasnat, M.M.; Hasan, S. Identifying tourists and analyzing spatial patterns of their destinations from location-based social media data. Transp. Res. Part C Emerg. Technol. 2018, 96, 38–54.

[10] Fukuda, D.; Kobayashi, H.; Nakanishi, W.; Suga, Y.; Sriroongvikrai, K.; Choocharukul, K. Estimation of Paratransit Passenger Boarding / Alighting Locations Using Wi-Fi based Monitoring: Results of Field Testing in Krabi City, Thailand. J. East. Asia Soc. Transp. Stud. 2017, 12, 2151–2169.

[11] Önder, I.; Koerbitz, W.; Hubmann-Haidvogel, A. Tracing Tourists by Their Digital Footprints: The Case of Austria. J. Travel Res. 2014, 55, 566–573.

[12] Musa, A.B.M.; Eriksson, J. Tracking unmodified smartphones using wi-fi monitors. In Proceedings of the Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems - SenSys ’12; 2012; pp. 281–294.

[13] Tsubota, T.; Yoshii, T. An Analysis of the Detection Probability of MAC Address from a Moving Bluetooth Device. Transp. Res. Procedia 2017, 21, 251–256.

[14] Danalet, A.; Tinguely, L.; Lapparent, M. de; Bierlaire, M. Location choice with longitudinal WiFi data. J. Choice Model. 2016, 18, 1–17.

[15] Potortì, F.; Crivello, A.; Girolami, M.; Barsocchi, P.; Traficante, E. Localising crowds through Wi-Fi probes. Ad Hoc Networks 2018, 75–76, 87–97.

[16] Abedi, N.; Bhaskar, A.; Chung, E. Tracking spatio-temporal movement of human in terms of space utilization using Media-Access-Control address data. Appl. Geogr. 2014, 51, 72–81.

[17] Dunlap, M.; Li, Z.; Henrickson, K.; Wang, Y. Estimation of Origin and Destination Information from Bluetooth and Wi-Fi Sensing for Transit. Transp. Res. Rec. J. Transp. Res. Board 2016, 2595, 11–17.

[18] Kusakabe, T.; Yaginuma, H.; Fukuda, D. Estimation of bus passengers’ waiting time at a coach terminal with Wi-Fi MAC addresses. Transp. Res. Procedia 2018, 32, 62–68.

[19] Fournier-Viger, P.; Chun-Wei Lin, J.; Kiran, R.U.; Koh, Y.S.; Thomas, R. A Survey of Sequential Pattern Mining. Data Sci. Pattern Recognit. 2017, 1, 54–77.

[20] Pitchayadejanant, K.; Nakpathom, P. Data mining approach for arranging and clustering the agro-tourism activities in orchard. Kasetsart J. Soc. Sci. 2018, 39, 407–413.

[21] Wang, F.; Li, K.; Duić, N.; Mi, Z.; Hodge, B.M.; Shafie-khah, M.; Catalão, J.P.S. Association rule mining based quantitative analysis approach of household characteristics impacts on residential electricity consumption patterns. Energy Convers. Manag. 2018, 171, 839–854.

[22] Gopikaramanan, R.; Rameshkumar, T.; Senthil Kumaran, B.; Ilangovan, G. Novel control methodology for H-bridge cascaded multi-level converter using predictive control methodology. Glob. J. Pure Appl. Math. 2015, 11, 3053–3071.

[23] Park, S.H.; Jang, S.Y.; Kim, H.; Lee, S.W. An association rule mining-based framework for understanding lifestyle risk behaviors. PLoS One 2014, 9.

[24] Tan, P.-N.; Kumar, V.; Srivastava, J. Selecting the right objective measure for association analysis. Inf. Syst. 2004, 29, 293–313.

[25] Tank, D.M. Improved Apriori Algorithm for Mining Association Rules. Int. J. Inf. Technol. Comput. Sci. 2014, 6, 15–23.

[26] North, M. Data Mining for the Masses; 2012; ISBN 9780615684376.

[27] Shou, Z.; Di, X. Similarity analysis of frequent sequential activity pattern mining. Transp. Res. Part C Emerg. Technol. 2018, 96, 122–143.

[28] García, E.; Romero, C.; Ventura, S.; Calders, T. Drawbacks and solutions of applying association rule mining in learning management systems. In Proceedings of the Proceedings of the International Workshop on Applying Data Mining in e-Learning (ADML 2007); Crete, Greece, 2007; pp. 13–22.

Chapter 3

Association Rule Mining Tourist-Attractive Destinations for the Sustainable Development of a Large Tourism Area in Hokkaido Using Wi-Fi Tracking Data

3.1 Introduction

Investigating tourist mobility, including trips to urban tourism destinations, is of widespread interest in tourism management [1]. Several factors have contributed to the growth of urban tourism: Changing work/leisure patterns, such as traveling on long weekends, has led to more frequent short holidays; there is an overall improvement in mobility; low-fare tickets have become available; and there is a pervasive desire for new experiences [2]. Numerous attempts have been made to measure the competitiveness of cities, regions or countries as tourist destinations and identify the factors that contribute to their enhanced and sustained competitive positions [3]. Tourist movement within a destination is often regarded as a black box that characterizes the behavior of an individual tourist rather than an aspect that should be explored and explained [4]. However, the movements of travelers within large tourism areas are important to understand because they play a fundamental role in destination management strategies, including route and activity planning, tourism products, attraction planning, and accommodation development.

The benefit of expanding the current understanding of traveler movements for an individual tourist spot involves evaluating and determining whether information provided about a specific location can be used to improve traveler experiences [5]. Within Hokkaido, the northern region of Japan, the tourist population has risen significantly in recent years. The inbound traveler population size has increased from 46 million in 2011 to 55 million in 2017, and it is continually increasing. Domestic travelers represent the largest portion of tourists in this area, although international travelers have also been increasing in recent years [6]. Rich natural scenery, high-quality ecotourism management, and friendly tourism environments have made Hokkaido a destination that is frequently selected by tourists around the world. This area is well known for its volcanoes, natural hot springs, outdoor activities, and ski resorts. The region covers an area of 83,000 km^2 and encompasses unique terrains, alpines, and cities. However, distance and travel time are major factors that determine the accessibility of attractive destinations located in the countryside, and the use of private transportation often needs to be considered. To reach their desired destinations, travelers tend to choose private modes of mobility, such as private vehicles and rental services, rather than public transportation. The frequent decision to travel privately usually leads to traffic congestion, especially during the tourism season. Thus, sustainable tourism policies have the potential to maintain the tourist area’s atmosphere and develop tourism sustainability.

In recent years, numerous probing studies [7-11] have been conducted on tourism and visitor destinations. However, data collection has presented several constraints, such as survey expense, the bias of questionnaire results, and the accuracy of the measurement devices used for observation. It is thus challenging to implement such studies, especially in large areas or over a long observation period. Consequently, to understand the phenomenon of traveler movement in a large tourism area, the use of modern detection technology, such as Wi-Fi scanner devices, has become widespread and is revolutionizing the methods by which researchers study the mobility of people [12]. Innovative technologies for conducting surveys are introducing a new era of data collection and superseding conventional survey methods, such as traditional paper-based or call-based questionnaire methods and email-based interviews. These new technologies are capable of capturing high-resolution data on the behavior of individual travelers in a large-scale population [7]. These tools also provide greater accuracy and reliability than conventional methods. With the emergence of information, communication, and technology (ICT), smart technology has been able to mitigate the challenges of data collection and thus change the approach to survey studies [13,14].

In tourism studies, the appropriate data mining technique is essential for detecting trajectory patterns in a complex dataset. Association rule mining (ARM), also as known as market basket analysis (MBA) in business research, is a type of machine learning analysis used as a data mining technique for discovering relationships between variables in extensive datasets. It is an essential and popular method in the business field, including medical and pharmaceutical research [15], as well as transport and tourism studies. Tourist planners can identify relations between one tourism spot and another in order to manage or propose planning for tourism development. Association rule learning has been adopted in tourism management for mining association rules between locations visited by tourists [16]. Articles have explored associations among different visitor segments; they have also suggested the potential use of tracking devices to observe tourist behavior. Several mining techniques have been applied to explain the traveling behavior of tourists. These techniques consist of classification, prediction, and association.

This research aims to examine the digital footprint of travelers in the massive tourism area of Hokkaido. The study has three contributions to the field of tourism research. First, a new survey method was implemented to study tourism: A Wi-Fi scanner device was used to capture radio transmissions from travelers' Wi-Fi-enabled devices. Probe data were generated on the basis of probe request data, which contain specific details, such as media access control (MAC) addresses, dates and times of travel, and the sites visited. Second, market basket analysis was applied to extract the association rules between tourism destinations, and the rules are visualized to facilitate understanding. Third, in the context of destination management, a conceptual discussion on the results of the data mining procedure is provided, and a sustainable tourism approach is suggested in the conclusion.

3.2 Literature review

The area of tourism research involves various scientific principles, including management and development. For destination management in a massive tourism area, appropriate destinations that are positively experienced by travelers should be identified; this study attempts to meet this objective using a new survey method. Thus, this section points out several essential elements in this approach—digital footprints in tourism studies, market basket analysis in tourism studies, destination management and sustainability, and destination types and competitiveness.

3.2.1 Digital Footprint in Tourism Studies

Digital footprints [8] are widely used in social mobility studies, and tracking technology and data extraction processes are acknowledged as integral in overcoming previous challenges in data collection. Therefore, social media platforms are considered to be human probe data sources, while radiofrequency detectors serve as the means of extracting digital footprints from wireless communications. According to social media tracing [17], various past studies have used social media platforms, such as Twitter and Flickr, to explore traveler behavior and trajectory patterns, with a focus on smartphone users in wide tourism areas [7,8].

Recently, the exploration of traveler mobility using smart sensing technologies has become more reliable [18,19]. Smart sensors capture anonymous, specific identifiers from each wireless communication device within scanning range. A MAC address is unique to an individual device, and it gives direct information on the movement of the enabled device along covered routes [20]. For example, from a business viewpoint, the market shares of a new catering location can be predicted by using Wi-Fi tracking to detect the sequences of activities occurring in a specific area, and the MAC address and username capability enhance the model's location choice [21]. Plug-and-play Wi-Fi sensors [22] equipped with USB ports have been configured to collect Wi-Fi signals with the aim of detecting human presence at a specific indoor location. Abedi's study highlighted mobility tracking as a key benefit of the MAC address; similarly, it can be used to track the spatiotemporal movement of an individual in terms of space utilization [23].

According to related articles, which mainly address small spaces such as a room in a campus building or within the campus area, few studies have observed larger areas. Thus, case studies applied to more extensive areas are being considered; for example, origin and destination data were used to estimate city transit mobility in a large area [9]. Moreover, the authors in [13] investigated paratransit passenger boarding and deboarding locations in a specific tourism area by using a Wi-Fi scanner device based on Raspberry PI: Its advantages include an inexpensive unit cost and ability to be implemented in several places. Wi-Fi probe requests [24] were utilized to estimate the waiting times of bus passengers at a terminal from the footprints of travelers' Wi-Fi-connected devices, such as smartphones. In different studies, the MAC address has served as a digital footprint, and this unique information can be applied to observe people's movements. The passive Wi-Fi tracking process to acquire digital footprints has been applied to detect the location of a single phone and provide a large-scale spatiotemporal trajectory, similar to GPS tracking [2] for detected wireless communication devices [19].

3.2.2 Market Basket Analysis in Tourism Study

In tourism studies, the movement of travelers between destinations and their spatial relationship in a large area can be complex [25]. Several previous studies in transport and travel sciences have extracted information from massive trajectory databases, and they have tended to use sequential pattern mining [17]. Alternative data mining methods include association rule mining analysis [10,26-28], which is a rule-based machine learning method. It has been used to identify groups of variables that are highly correlated with each other from an extensive database or determine patterns of relations between variables of interest [11]. Besides ARM, market basket analysis [29] has been widely used in retail businesses and marketing research for many decades, and it allows retailers to identify relationships between purchased items and buyer behavior. The most popular algorithm is Apriori, which has been used to extract the frequency of itemsets from massive databases and determine association rules to obtain the desired information [30,31]. This type of analysis works by looking for items that frequently occur together in transactions.

Moreover, to understand traveler behavior, pattern mining techniques are suitable for detecting hidden patterns in extensive databases, which is relevant to tourism and destination management research [32]. An example of applying ARM to tourism is the integration of the ARM technique with Bluetooth tracking data to identify mobility patterns in a tourist attraction area [16]. The ARM data mining technique was also applied to establish association rules to find patterns of activities in orchard tourism [33]. Furthermore, a case study on Hongkong presented a new approach that extends the capability of the association rule technique in an aim to capture changes and trends in outbound tourism [11]. Therefore, these methods present an opportunity to apply the market basket concept to tourism research for tourism planning and management [34].

3.2.3 Destination Management and Sustainability

Destinations are a blend of tourism services and experiences [35]. The theoretical concepts underlying mass tourism destination studies are a combination of five measurable dimensions— geographical, temporal, compositional, social, and dynamic dimensions [36-38]. The specific characteristics of an attraction area can affect management policies or even development planning. Tourism destinations that are regarded as attractive are defined by geographical areas, such as flower gardens, mountains, rivers or even cities. The massive destination concept, which interprets or judges a visitor as a unique entity, has been increasingly recognized. Thus, a massive tourism area requires a destination management organization (DMO) to be accountable for marketing and planning destination resources, and DMOs are essential for destination management and sustainable tourism [35,37]. Because of the sharing of resources, such as flower gardens that are shared by various stakeholders, tourist destinations are perceived and experienced as part of the same area by visitors. One article [35] proposed a framework of six components that are necessary for tourism destination analysis: Attractions, accessibility, amenities, available packages, activities, and ancillary, which are a combination of products, services, and experiences that are provided locally. These aspects can generate satisfaction among interacting tourists and local business owners when a DMO emphasizes total management rather than merely marketing. Moreover, this framework is useful for monitoring tourist satisfaction levels, and its components can be used as criteria for the success of tourism management and development [38]. On the other hand, small-scale development may only permit destination development on low-grade land, or it may be considered low grade from scenic and other viewpoints [39].

Furthermore, emphasizing today's greater public involvement in destination planning is essential when planning for tourist development. Another essential factor in the planning and marketing process is the destination scale, which is considered to be a critical part of destination development that satisfies tourist objectives [40]. In the context of sustainability and ecotourism development, the authors of Reference [41] examined changes that have taken place in politics, policy, development, conservation, human–environmental relations, and the convergence of these areas. They also proposed seven preliminary steps toward a greater understanding of sustainable tourism and recommended adding post-normal science to eliminate current and ineffective methods for studying tourism [41]. Extrapolative methods, such as time series forecasting models, have been highly successful in tourism science, and they can obtain several data patterns for structuring models. Similarly, historical series data, such as seasonal patterns or seasonality, are generally used in the design phase of tourism and necessitate the forecasting process, which includes factors such as climate, holidays, business customs, and calendar effects [42].

3.2.4 Types of Destination and Competitiveness

Travel behavior classification and segmentation research has become increasingly more complicated as modern travelers combine pleasure with business to gain time and cost advantages. These two major categories are reasonably identifiable, and they are treated differently in this text to simplify the concepts and facilitate the marketing response [35]. Moreover, destinations should be aware of not only the current demand but also potential markets they can attract. Business trips and leisure visitors have some specific characteristics related to the alignment between the type of destinations visited and activities undertaken, as shown in Table 3.1.

Table 3.1 Type of destinations, ideal customer, and activities undertaken [35].

Type of destination	Target market	Activities
CBD[1], Urban	Business	MICE, Education, religion, health
	Leisure	Sightseeing, shopping, shows, short breaks
Seaside, Ocean	Business	MICE
	Leisure	Sea, sun, sand, sex, sports
Alpine, Mountain	Business	MICE
	Leisure	Activities, sports, health
Rural, Regional	Business	MICE
	Leisure	Relaxation, agriculture, activities, sports, health
Authentic third world	Business	Exploring business opportunities, incentives
	Leisure	Adventure, authentic, special interest
Unique, Exclusive	Business	Meetings, incentives, retreats
	Leisure	Special occasion, honeymoon, anniversary

[1] Central business district.

The point of view of this table illustrates what the business market often refers to as meeting-incentives-conferences-exhibitions (MICE). Hence, the competitiveness in the tourism industry may be an inevitable consequence of such models, which may be considered fundamental elements of comparative and competitive advantages [37]. As new tourism competition emerges [43], the industry demands new perspectives on consumers, technologies, production practices, and management techniques. It has been proposed that destination competitiveness aims to achieve certain goals in various relevant areas, which are also categorized into particular elements—economic, attractiveness and satisfaction, and sustainability [3]. Furthermore, destination competitiveness will become a fundamental principle of the competitive tools that control the process of value creation in the tourism industry, and it is essential for understanding the relations and interactions between factors of competitiveness [43].

3.3 Methodology

3.3.1 Tracking Equipment

The scanner consists of a Raspberry Pi 2, which is a hardware operating system unit. This equipment has been developed to detect Wi-Fi transmissions and is capable of automatically capturing probe requests, which are emitted continuously from Wi-Fi-connected devices, such as smartphones. The operating range of coverage has a radius of about 400 meters under ideal conditions [13]. The scanner is integrated with several components—an external USB antenna, powered by a portable battery pack or standard USB charging cable, and an SD memory card. The Wi-Fi scanner is enhanced by software that is equipped to enable the device to collect probe requests. It captures individual MAC addresses and several spatial–temporal-specific data, which are then converted into the simplified format shown in Figure 3.1. The probe request intervals depend entirely on the individual emitting source, and the communication function enables data to be transmitted to the server at five-minute intervals (some sensors equipped with a Wi-Fi router are capable of real-time transmission to the server).

mac	day	time	site	sitename
00009872164fb495849d2de...	712	15:17:24	28	Shikisainooka
00009872164fb495849d2de...	712	15:53:26	28	Shikisainooka
00009872164fb495849d2de...	712	17:09:31	25	FuranoMarche
00009872164fb495849d2de...	712	17:22:30	25	FuranoMarche
00009872164fb495849d2de...	712	17:39:32	25	FuranoMarche
00009872164fb495849d2de...	712	17:44:33	25	FuranoMarche
00009872164fb495849d2de...	712	19:08:36	10	PAIwamizawaDown
00009872164fb495849d2de...	712	19:19:36	10	PAIwamizawaDown

Figure 3.1 Example of probe requests data from Wi-Fi scanner devices.

3.3.2 Data Collections and Destinations

This research was conducted on specific target areas, and it involved traveler mobilities and activities in a massive region, namely, the central and northern regions of Hokkaido, as shown in Figure 3.2. Furano and neighboring cities have been recognized as flower garden areas that attract massive international and domestic tourists each year, especially during the lavender blooming season in June and July. Thus, the investigation period was from June 19 to July 23, 2017. Multiple Wi-Fi sensors were equipped in 31 different observation destinations. Besides this setup, multiple devices were placed in single areas in various facilities (e.g., buildings, farms, or gardens) located in the same area. The field experiment was conducted at traffic attraction nodes, such as the service and parking area (SA/PA) on a road network, as well as tourist attraction destinations, such as hotels, the zoo (AsahiyamaZoo), famous flower gardens, car rental shops, and recreation and shopping areas, as shown in Table 3.2.

3.3.3 Data Preparation

To perform the ARM analysis using R version 1.1.463, dataset preparation was required. This study entailed detecting patterns of transactions from the probe request dataset, and only continuous transactions occurring over time were considered to be transaction data (see Figure 1). Figure 1 shows an example of transaction records and a continuous pattern that is appropriate for ARM. Since the same MAC address is recognized, and the location in the last column identifies different places, these characteristics, along with the date and time, were selected for the dataset. Because the MAC number identifies the person who visited the location and never changes, if the sensors capture the same MAC number transmitted in different locations, it represents the movement of the traveler from one location to another.

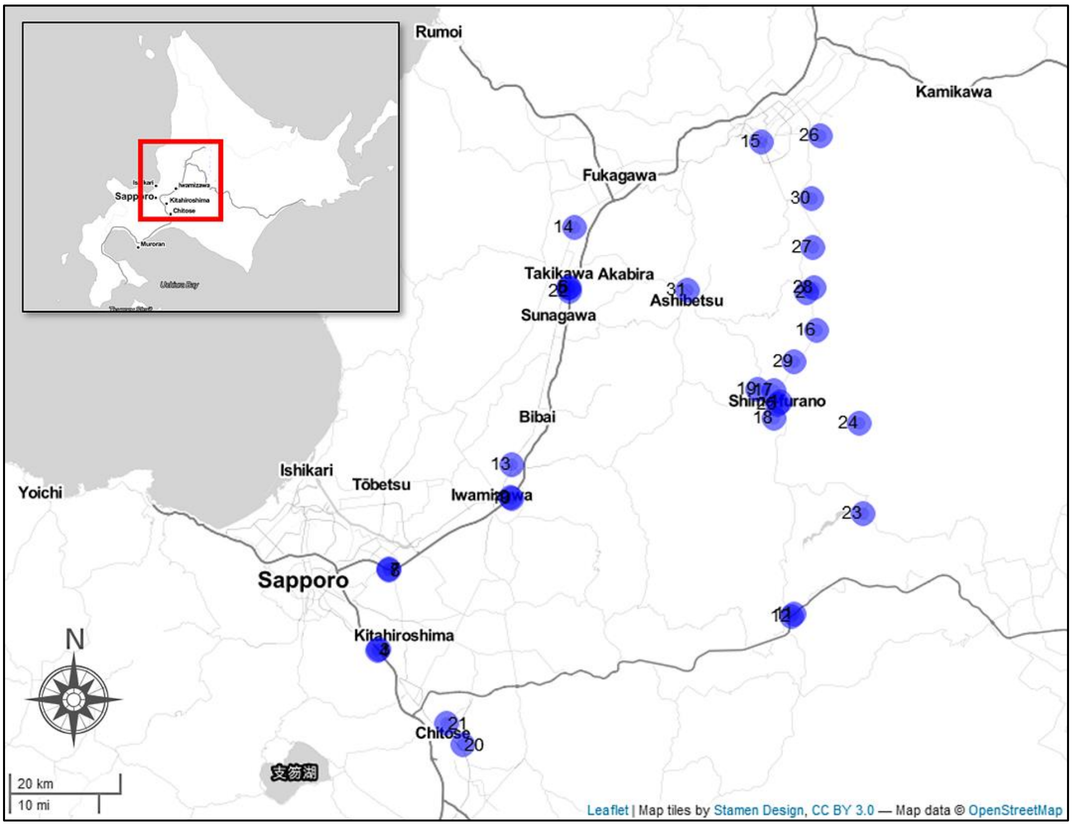


Figure 3.2 The coverage of observation locations of attractive destinations in Hokkaido.

Table 3.2 The list of observation locations.

ID	Type	Location	ID	Type	Location
D1	AC	NaturaxHotel	D17	RE	CampanaRokkatei
D2	FG	LavenderOwnerGarden	D18	RE	FuranoCheesePlant
D3	PA	PAWattsuUp	D19	FG	LavenderHighlandFurano
D4	PA	PAWattsuDown	D20	CR	ShinchitoseTOYOTA
D5	PA	PASunagawaUp	D21	RS	RSSalmonpark
D6	PA	PASunagawaDown	D22	SA	SAHighwayoasis
D7	PA	PANopporoUp	D23	RS	RSMinamifurano
D8	PA	PANopporoDown	D24	RE	Gorohouse
D9	PA	PAIwamizawaUp	D25	RE	FuranoMarche
D10	PA	PAIwamizawaDown	D26	RE	AsahiyamaZoo
D11	PA	PAShimukappuUp	D27	RS	RSBieinokura
D12	PA	PAShimukappuDown	D28	FG	Shikisainooka
D13	RS	RSMikasa	D29	FG	NakafuranoLavenderGarden
D14	RS	RSTakikawa	D30	CR	AsahikawaTOYOTA
D15	RS	RSAsahikawa	D31	RS	RSStarPlazaAshibetsu
D16	FG	FlowerLandKamifurano			

AC = accommodation, FG = flower garden, PA = parking area, RS = rest area, CR = car rental shop, SA = service area, RE = recreation destination.

3.3.4 Destination Rule Mining

Association rules form an essential class of regularities in data, and determining these rules is a fundamental data mining process by which all co-occurrence relationships among tourism destination data are found [16,44]. However, to obtain the best rules from ARM, the interestingness of association patterns were evaluated by following a framework of principles, i.e., the support, confidence, and lift of a rule [45].

The support of a rule indicates how frequently the group of items appear together. It is a useful measure since a high value of support represents a repeated frequency, which means that the itemset is often detected [31]. The support is calculated as expressed in Equation 3.1. The confidence of a rule represents the accuracy or conditional probability of the itemset appearing, and it is given in Equation 3.2. Its value is related to a set of transactions: That which contains X also contains Y. An association rule with a high confidence value is considered reliable.

$$Support\ (X \Rightarrow Y) = \frac{Number\ of\ transactions\ with\ both\ X\ and\ Y}{Total\ number\ of\ transactions} = P(X \cap Y) \tag{3.1}$$

$$Confidence(X \Rightarrow Y) = \frac{Number\ of\ transactions\ with\ both\ X\ and\ Y}{Total\ number\ of\ transactions\ X} = \frac{P(X \cap Y)}{P(X)} \tag{3.2}$$

Nevertheless, the confidence itself may not be enough to explain the result; thus, the lift value is applied [46]. The lift of a rule expresses the general measurement of the association between the itemset. It is defined by the confidence of the rule divided by the expected confidence, assuming that the itemsets are independent; the expected confidence and lift are defined in Equations 3.3 and 3.4, respectively. A lift value equal to 1 specifies zero correlation; a lift greater than 1 specifies a positive correlation; and a lift less than 1 specifies a negative correlation [28]. Thus, higher lift ratios indicate stronger associations, and it reflects the true relation of an itemset beyond what is expected to be found by chance [33].

$$Expected\ Confidence(X \Rightarrow Y) = \frac{Number\ of\ transactions\ with\ Y}{Total\ number\ of\ transactions} = P(Y) \tag{3.3}$$

$$Lift(X \Rightarrow Y) = \frac{Confidence(X \Rightarrow Y)}{Expected\ Confidence(X \Rightarrow Y)} = \frac{P(X \cap Y)}{P(X).P(Y)} \tag{3.4}$$

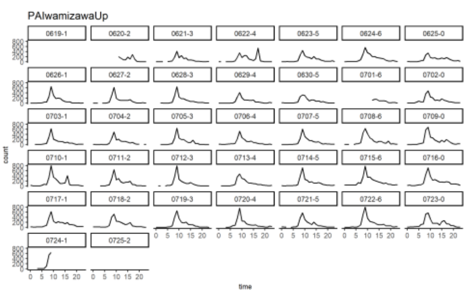
The ARM method has been used not only for market research but also for tourism, urban development, and various other fields of study. For example, ARM has been applied in areas such as sequential pattern mining analysis to extract travel patterns in transportation research [17]; it has also been used for series pattern mining as correlation rule analysis in big data analytics [47]. The method’s mechanism is based on if-then statements concerning the variables, and it then identifies the association of an interesting itemset [48]. These are the advantages of ARM in destination mining development. It is also possible to visualize the combination of location order and travel patterns. Furthermore, to visualize the rules of destination analysis, the package arulesViz, which is an interactive visualization technique in the R package [49], was applied. It enhances the analysis by using color shading, reordering, and interactive features [50] that clarify the destination perspective and visitor behaviors in a massive tourism area.

3.4 Results

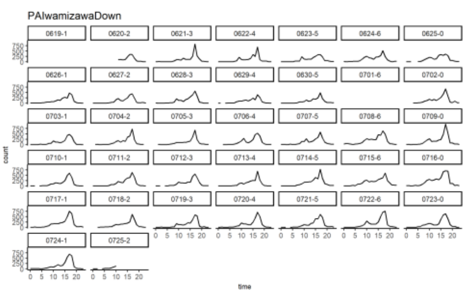
3.4.1 Hourly Patterns in Time Series

Since Wi-Fi tracking data were used to analyze visitors' mobility in the study area, time series analysis was deemed a suitable method to clarify mobility patterns, which are expressed as the hourly and weekly demand at a particular attraction. The hourly demands at parking and service areas along the expressway network and at tourism destinations, both inbound and outbound directions, are illustrated in Figure 3.3(a) and 3.3(b). The scanner devices captured the trace of visitors who spent time in the parking and service stations. The patterns show the visiting occurrence during the day between early morning and midnight, as public facilities and services are provided 24/7. For example, the inbound direction "PAIwamizaUp" rush hour appears during the morning for the starting stage of travelers' journeys, while the outbound "PAIwamizawaDown" rush hour generates the reverse pattern because visitors are heading home at the end of the day.

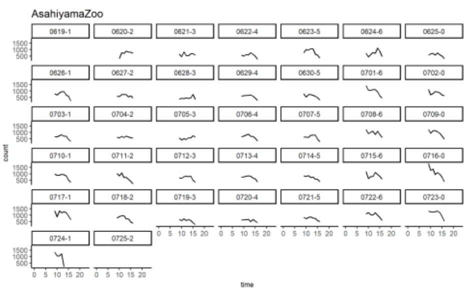
For attractive tourist destinations, such as the zoo "AsahiyamaZoo" and flower garden "Shikisainooka", the illustration shows that travelers move the most during the daytime, which reflects the operating time of the destinations, as shown in Figure 3.3(c) and 3.3(d). With more detailed time-zone aggregation, it can be confirmed that the visiting peaks occur from 10 a.m. to 3 p.m. during the day and decline thereafter until 6 p.m. Furthermore, the trend changes according to several factors, such as traffic conditions, disasters, weather, and so on [42], which is a potential topic for a future study. The trends also reveal a one-day trip for the clustering destination analysis more details in Appendix A1.



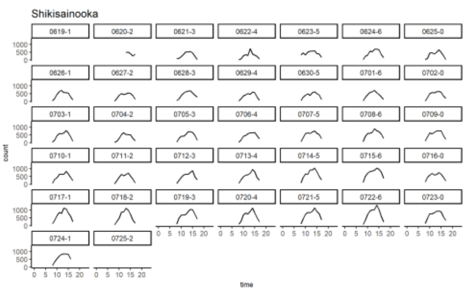
(a)



(b)



(c)



(d)

Figure 3.3 Hourly patterns of service and parking stations: (a) Inbound; (b) outbound. Hourly patterns at attractive tourist destinations: (c) The zoo; (d) flower garden.

3.4.2 Weekly Patterns in Time Series

The weekly patterns of visits to attractive destinations are represented by a collection of destinations as an example. The outcomes of monitoring “SAHighwayoasis” and “AsahiyamaZoo” represent trends in the number of visitors in each week, including the weekly peak, during the months of observation. The figures show that the weekend generates a greater number of visitors than a weekday; also, the trends for a service and parking area see Figure 3.4 and an attractive destination see Figure 3.5 reveal similar results, with a both showing a significant increase in visitors on weekends. Moreover, the traveler population increases closer to the summer, which is likely because it is the best period for lavender flower sightseeing and other attractive activities. Thus, the difference between weekday and weekend scenarios must be considered, as the number of visitors fluctuates over time Appendix A2.

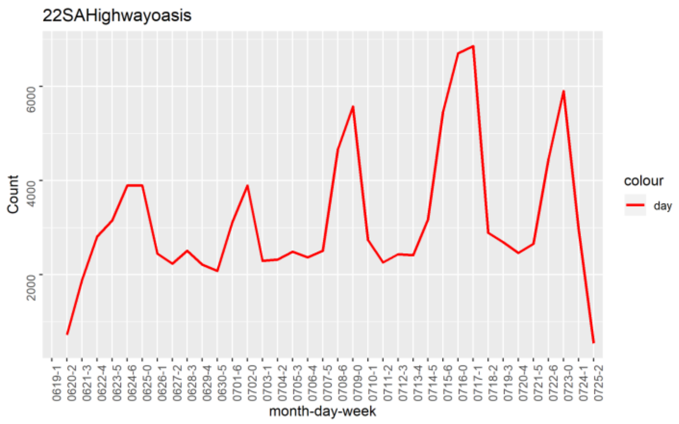


Figure 3.4 Weekly patterns of service and parking area at SAHighwayoasis station.

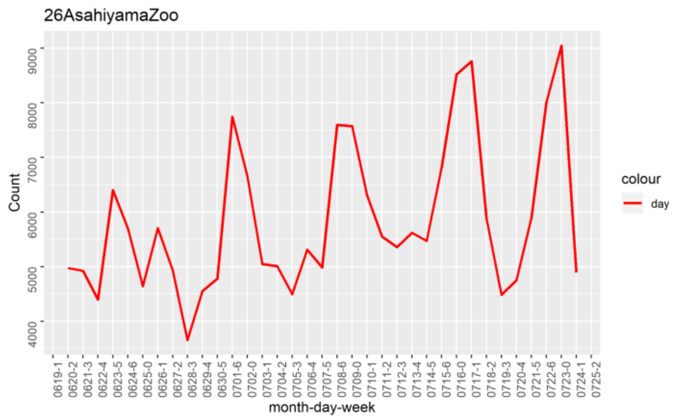
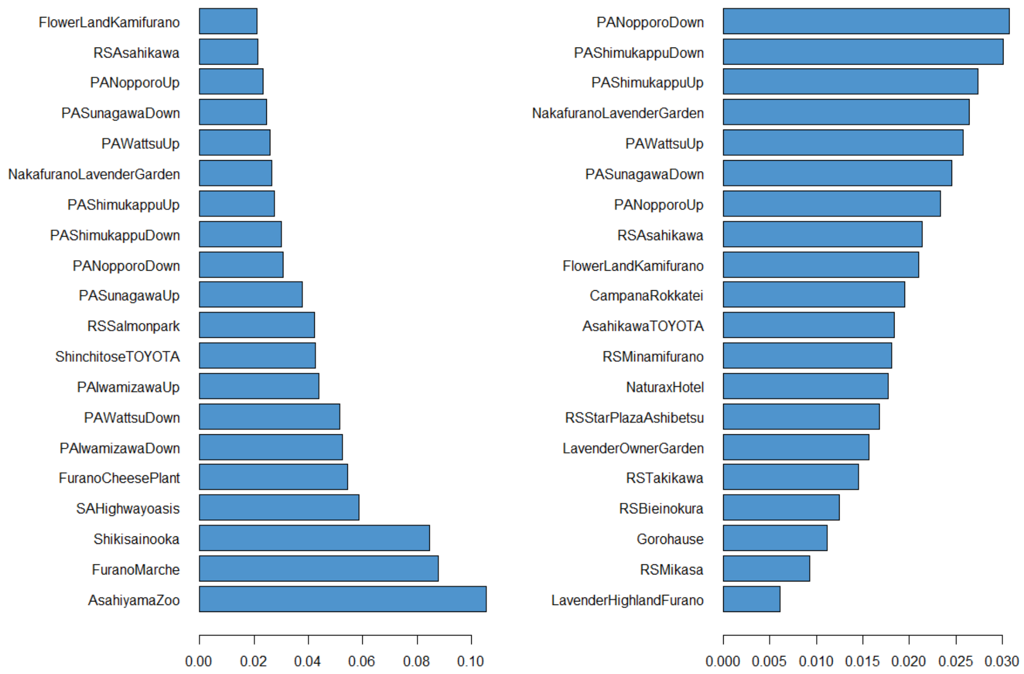


Figure 3.5 Weekly patterns of the tourism destination at AsahiyamaZoo.

3.4.3 Assessing of Attractive Destinations

Mining the most attractive destination was performed by probe requests, and the dataset in this study includes 1,048,575 records and four attributes: MAC number, day, time, and site name. This section reports the mining of famous destination visits. The distribution of the transaction indicates the frequency of visitations to the destination, and the relative distributions in Figure 3.6 reveal the most popular and the least visited destinations. The lift bar chart shows the top 20 most visited sites, among which “AsahiyamaZoo” was the most visited, showing the highest visitor population; this is a reasonable finding since this is the most famous zoo in Hokkaido. Conversely, “LavenderHighlandFurano” was the least visited destination.



(a) (b)

Figure 3.6 Relative distribution of visiting frequency

(a) 20 most visited sites; (b) 20 least visited sites.

Moreover, the plot shows that “AsahiyamaZoo”, “FuranoMarche”, and “Shikisainooka” had the most visitors when compared with other destinations. Hence, to improve the number of visitors to “FuranoChessePlant” or “FlowerLandKamifurano”, the planner or local tourism authorities could consider addressing the information related to the top destinations. Additionally, the trend bar charts will shift as a result of seasonality tourism [42]; for example, during winter, alpine ski resorts are likely to be more popular destinations than flower gardens.

3.4.4 Association Rule of Attractive Destinations

This section discusses the identification of association rules found in the huge dataset using Apriori algorithm-based ARM. Accurate results require the completion of tasks, such as finding association rules and identifying and removing redundant rules. Identifying redundant rules is an essential process to screen for duplicated rules, which must be filtered out for the analysis of appropriate rules [51].

The transactional database table shows an example of an itemset. The third transaction is an itemset with a length of two items, which are representative of the destinations visited by this tourist during his or her trip see Table 3.3. The distribution of rule lengths indicates visitors' preferences for how many locations they visit on their trip. Table 3.4 shows that 440,061 transactions were subjected to ARM processes for the extraction of the lengths of visiting destination rules. Most of the travelers visited one or two places on their journey, and the computation of a length of three indicators or more was less common. In order to extract meaningful transactions, thresholds were set for the degree of support, confidence, and lift. Then, excursion pattern tendencies were visualized. Since the number of transactions is large, the degree of support tends to be small, and the market basket principle was adapted to carry out destination analysis. Hence, for the ARM process, a confidence threshold of 1% was set as the minimum acceptable ratio, and the outcome must greater than or equal to the minimum ratio to meet the requirement for analysis. In the ARM process, the antecedent is the origin of the transaction, and the consequent is the final observation area. After removing redundant rules, only 13 rules remained, all of which have a rule length is contain two destinations. The remaining rules are sorted by decreasing confidence in Table 3.5.

Table 3.3 Example of transactional database for destination rule analysis.

Transaction ID	Length	Items (Destinations)
1	2	Shikisainooka , SAHighwayoasis
2	1	AsahiyamaZoo
3	2	FuranoMarche, FuranoCheesePlant
4	3	PANopporoUp, FuranoMarche, RSStarPlazaAshibetsu
:	:	:
440,061	3	PANopporoUp, FuranoMarche, NaturaxHotel

Table 3.4 The distribution of rule lengths.

Length	1	2	3	4	5	Total
Transaction	417,695	22,350	11	4	1	440,061

The rules present a pattern of destination visits, and the analysis considered transactions including two or more rules, i.e., rules with a length of one were excluded. The confidence threshold shows that about 6% of those who visited the parking area “PAIwamizawaUp” were highly likely to visit “PAIwamizawaDown”. Similarly, about 3% of those who visited “FlowerLandKamifurano” also visited “Shikaisainooka”, where there is a free-of-charge flower garden in the Furano area. Once the minimum confidence ratio is reached, it becomes a matter of focusing on rules with high confidence to identify the relationship between rules and destinations for development purposes. “Shikisainooka” seems to be a popular destination in this area, so it is relevant to find how it associates with other destinations, such as tourist and activity sites or even parking and service areas.

Table 3.5 The association rules of attractive destinations.

Rules ID	Association Rules	Support	Confidence	Lift
D9 → D10	{PAIwamizawaUp} => {PAIwamizawaDown}	0.0025	0.0575	1.0973
D16 →D28	{FlowerLandKamifurano} => {Shikisainooka}	0.0011	0.0529	0.6252
D5 → D10	{PASunagawaUp} => {PAIwamizawaDown}	0.0015	0.0385	0.7357
D10 → D28	{PAIwamizawaDown} => {Shikisainooka}	0.0019	0.0358	0.4227
D9 → D28	{PAIwamizawaUp} => {Shikisainooka}	0.0012	0.0264	0.3121
D5 → D28	{PASunagawaUp} => {Shikisainooka}	0.0010	0.0253	0.2987
D10 → D26	{PAIwamizawaDown} => {AsahiyamaZoo}	0.0012	0.0228	0.2160
D18 → D25	{FuranoCheesePlant} => {FuranoMarche}	0.0012	0.0222	0.2532
D9 → D22	{PAIwamizawaUp} => {SAHighwayoasis}	0.0010	0.0217	0.3713
D22 → D26	{SAHighwayoasis} => {AsahiyamaZoo}	0.0011	0.0189	0.1789
D18 → D28	{FuranoCheesePlant} => {Shikisainooka}	0.0009	0.0170	0.2014
D28 → D25	{Shikisainooka} => {FuranoMarche}	0.0010	0.0117	0.1334
D28 → D26	{Shikisainooka} => {AsahiyamaZoo}	0.0009	0.0109	0.1032

Moreover, the parking area also has a high lift value: The rule whereby “PAIwamizawaUP” is followed by “PAIwamizawaDown” (D9 → D10) has a lift value larger than 1.0, indicating that the two are dependent and have a positive effect on each other. Conversely, for others with small values, the antecedent has a negative effect on the consequent destination. This could reflect that visitors more frequently visited sites with high lift values. In some of the rules involving the parking and service areas, if the antecedent increases, then the consequent increases as well. There are typical mobility patterns from Sapporo to the Furano area for sightseeing and activities. For example, “PAIwamizawa” is often a final place; it seems to be a stopover area when returning to

Sapporo after sightseeing in the Furano area. Thus, it could be assumed that this place has the potential to be an effective part of a tourism development plan: For instance, repeat visits could be encouraged by disseminating tourist information at this site to enhance interest in visiting other areas.

To discover potential locations that should be promoted in the future, visualization techniques are essential for illuminating relationships. Interactive visualizations have become one of the most prominent types of graph-based visualization, which is commonly reviewed and implemented using the arulesViz algorithm. Graph-based visualization reveals how rules connect specific items. It is a viable method for small sets of rules [49], and it was used to depict the ARM process for destination analysis. In the graph-based method used here, arrows point from items to rules, vertices indicate antecedent items, and an arrow that points from a rule to an item indicates a consequent (see Figure 3.7).

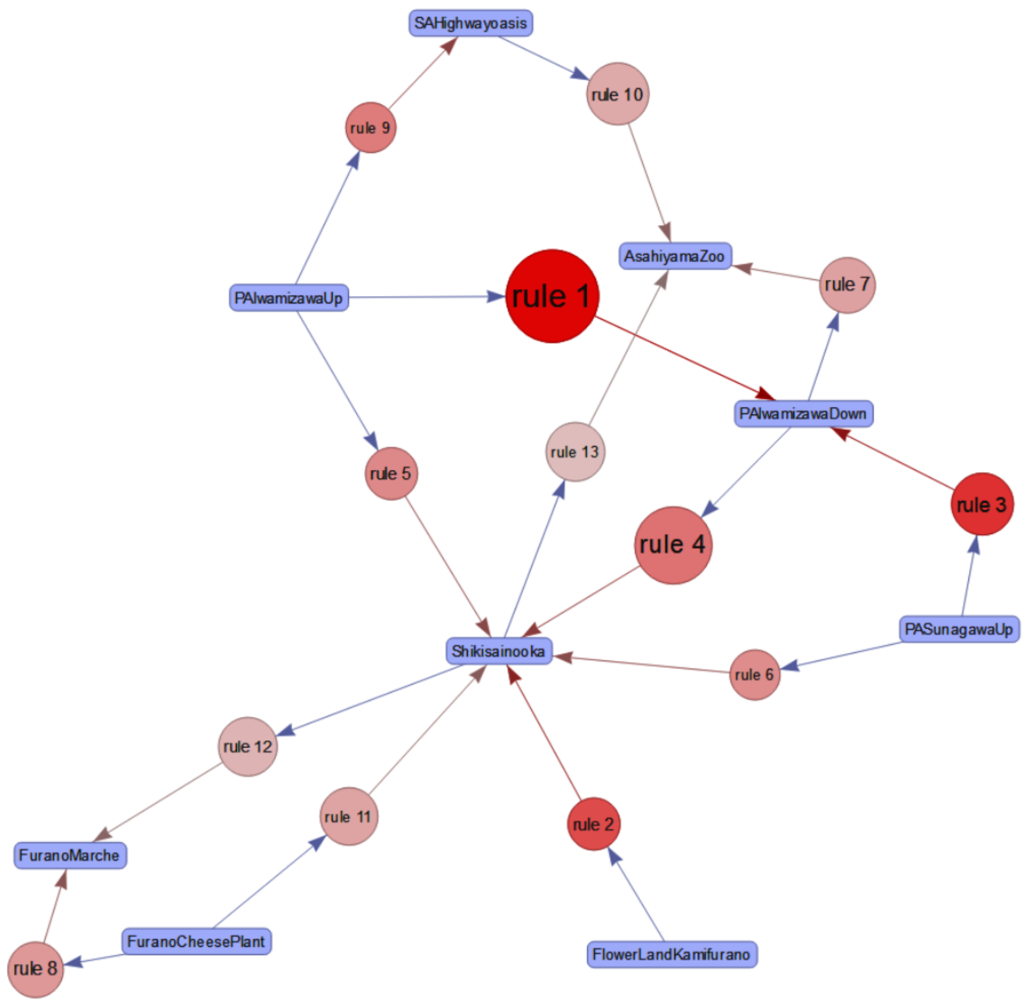


Figure 3.7 Visualization of all destination rules.

The plot shows arrows oriented toward “Shikisainooka”: According to ARM results, it is the most attractive location in the heartland of Furano. This flower garden attracted transactions from multiple locations, including garden and activity sites as well as the parking and service area. Similarly, “AsahiyamaZoo” is also an attractive destination. Further, a major contributor is “PAIwamizawaUp”, which generated the most rules in this area. It is recognized as an essential destination from which tourists disperse to other attractive locations, and it also has great confidence and lift. “PAIwamizawaDown”, a (SA/PA) station on the expressway, characterizes the intersection of the destination transaction. A well-represented rule describes visitors who visited “FlowerLandKamifurano” and then went to “Shikisainooka” (D16 → D28). Hence, this site may be considered a potential location to meet the demand of visitors in the future.

3.4.5 Implementation of Statistical Hypothesis Test

The finalized association rules in Table 3.5 reveal attractive tourism destinations, including service and parking areas, and they also reflect antecedent and consequent destinations. To examine the relationship between destinations, the Chi-squared test was applied to determine the statistical significance level of the association rules. The Chi-squared test is an analytical procedure that is used to measure the degree of dependence between variables [52]. Here, the results of this test are used to explain the relation between destinations and to discover whether one destination is dependent on another. In this analysis, the rules that contain only tourist attraction sites were selected for further analysis (i.e., parking and service areas were excluded; see Table 3.6). Hypotheses were established for the relationships between variables. In the Chi-squared test, a null hypothesis asserts that no relationship exists between an antecedent and a subsequent event, while the alternative hypothesis states that the relationship exists. Table 3.6 presents the statistical outcome for five destination rules.

For example, for the first rule, the Chi-squared statistic corresponds to a high significance level, with an alpha level of 0.1% and a p-value of 2.496E-4. This is strong evidence that the null hypothesis is rejected, thus supporting the existence of a relationship between “FlowerLandKamifurano” and “Shikisainooka”, represented by the rule ID (D16 → D28). Hence, the dependence experiment can be used to describe the degree of relation between destinations, and it is also a useful method for destination management and tourism policymaking.

Table 3.6 Chi-squared test for association analysis of tourism destinations ($p < 0.01$).

Rule ID	df	p-value	Null Hypothesis	Analysis
D16 → D28	2	2.496E-4	Reject	Association
D18 → D25	2	2.653E-3	Reject	Association
D18 → D28	2	2.928E-3	Reject	Association
D28 → D25	2	5.567E-3	Reject	Association
D28 → D26	2	7.175E-3	Reject	Association

3.4.6 Identification of Important Destinations

In the previous section, "Shikisainooka" is identified as a significantly essential destination in the study area and potentially beneficial for tourism management planning. Therefore, this section examines this location to determine whether it contributed to the massive tourism area of Hokkaido and the potential for expanding its scale so it can entertain the future demand of inbound and outbound visitors. To this end, ARM was enhanced to assess the contribution of this destination and identify factors that influenced visits to this remarkable site. Figure 3.8 and Figure 3.9 illustrate the inbound and outbound association rules resulting from ARM enhancement as defined in Tables 7 and 8. From the inbound rules, it is apparent that those who visited "FlowerLandKamifurano" frequently went on to visit "Shikisainooka" (D16 → D28); this rule has strong confidence compared with others. On the other hand, it can also be stated that before visiting "Shikisainooka", people tended to visit "FlowerLandKamifurano". Outbound rules show that after visiting "Shikisainooka", "PAIwamizawaDown" (D28 → D10) was often the next destination, and this rule also has strong confidence compared with others. Therefore, the results of this investigation provide significant insights that can be leveraged for tourism and destination management in this region.

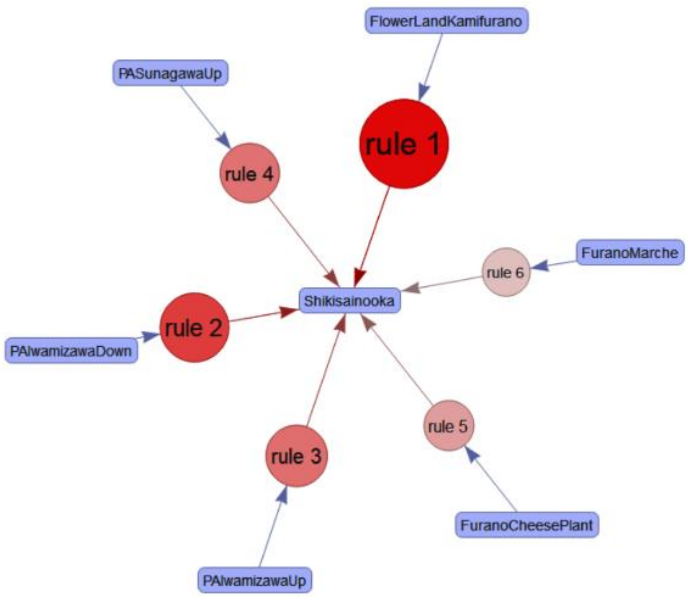


Figure 3.8 Recognized destination rules inbound visiting of Shikisainooka.

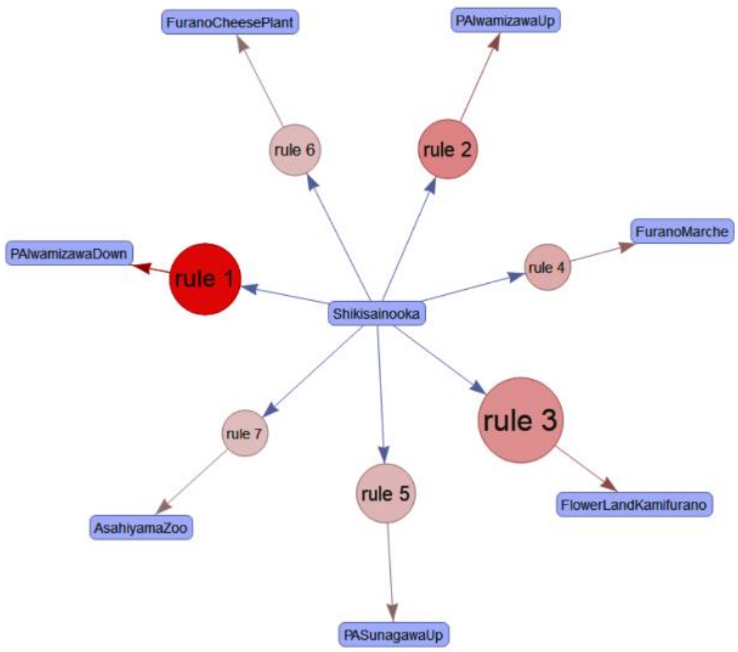


Figure 3.9 Recognized destination rules outbound visiting of Shikisainooka.

Table 3.7 Shikisainooka inbound transactions.

Rule ID	Antecedent	Consequent	Confidence	Lift
D16 → D28	FlowerLandKamifurano	Shikisainooka	0.0529	0.6252
D10 → D28	PAIwamizawaDown	Shikisainooka	0.0358	0.4227
D9 → D28	PAIwamizawaUp	Shikisainooka	0.0264	0.3121
D5 → D28	PASunagawaUp	Shikisainooka	0.0253	0.2987
D18 → D28	FuranoCheesePlant	Shikisainooka	0.0170	0.2014
D25 → D28	FuranoMarche	Shikisainooka	0.0113	0.1334

Table 3.8 Shikisainooka outbound transactions.

Rule ID	Antecedent	Consequent	Confidence	Lift
D28 → D10	Shikisainooka	PAIwamizawaDown	0.0221	0.4227
D28 → D9	Shikisainooka	PAIwamizawaUp	0.0137	0.3121
D28 → D16	Shikisainooka	FlowerLandKamifurano	0.0131	0.6252
D28 → D25	Shikisainooka	FuranoMarche	0.0117	0.1334
D28 → D5	Shikisainooka	PASunagawaUp	0.0113	0.2987
D28 → D18	Shikisainooka	FuranoCheesePlant	0.0109	0.2014
D28 → D26	Shikisainooka	AsahiyamaZoo	0.0109	0.1032

3.5 Conclusions

This study examined traveler movements among 31 destinations in Hokkaido to identify attractive tourist destinations and enhance sustainable tourism development in this area. Passive probe data from enabled wireless communication devices were integrated with the association rule mining method. The analysis involved the combination of several techniques, such as time series of visiting patterns, association mining implementation, and destination mining enhancement. These approaches were used to explore sequential patterns related to visiting attractive destinations. Moreover, the media access control number from enabled wireless communication devices was a key factor used to trace traveler movement in the targeted area. The application of association rule mining revealed the rules of destination visits. These rules depict strong relationships in terms of visit frequency, and they illustrate "if and then" statements for visitor movement analysis. These analytical methods are applicable to other tourist destination studies. The results show that, for weekend travel, factors such as seasonality tourism, distance to the location, and activities at the destination affect the type of destination selected and the number of destinations visited per trip. Moreover, visualization methods were used to translate mathematical logic to graphics. These finding could be useful for developing tourism strategies for destination management and planning in the Hokkaido region. The authors hope that this study provides novel information to the authorities and organizations involved in tourism management and allow them to make sound decisions on policies and plans to improve the tourism industry's environment and sustainable development.

Discussion on tracking device limitation: The advantages of using Wi-Fi tracking devices in this survey instead of conventional methods include their ability to capture a massive amount of data in the detection range, and they can cover large areas by installing multiple devices at different locations. The system of devices can be operated with few human resources in long-term operations and have benefits including powerless consumption and durability in various weather conditions. In this study, an offline device was used; this is a limitation since the amount of data acquired depends on the storage capacity and data resolution. A future study should consider an online device to enable real-time monitoring and eliminate the data storage problem.

3.6 References

[1] Sugimoto, K.; Ota, K.; Suzuki, S. Visitor Mobility and Spatial Structure in a Local Urban Tourism Destination: GPS Tracking and Network analysis. Sustainability 2019, 11, 919.

[2] Edwards, D.; Griffin, T. Understanding tourists' spatial behaviour: GPS tracking as an aid to sustainable destination management. J. Sustain. Tour. 2013, 21, 580–595.

[3] Abreu Novais, M.; Ruhanen, L.; Arcodia, C. Destination competitiveness: A phenomenographic study. Tour. Manag. 2018, 64, 324–334.

[4] Mckercher, B.; Lau, G. Movement Patterns of Tourists within a Destination. Tour. Geogr. 2008, 10, 355–374.

[5] East, D.; Osborne, P.; Kemp, S.; Woodfine, T. Combining GPS & survey data improves understanding of visitor behaviour. Tour. Manag. 2017, 61, 307–320.

[6] JNTO-Japan National Tourism Organisation Japan Tourism Statistics Available online: https://statistics.jnto.go.jp/en/graph/#category--6 (accessed on Jan 22, 2019).

[7] Hasnat, M.M.; Hasan, S. Identifying tourists and analyzing spatial patterns of their destinations from location-based social media data. Transp. Res. Part C Emerg. Technol. 2018, 96, 38–54.

[8] Önder, I.; Koerbitz, W.; Hubmann-Haidvogel, A. Tracing Tourists by Their Digital Footprints: The Case of Austria. J. Travel Res. 2014, 55, 566–573.

[9] Dunlap, M.; Li, Z.; Henrickson, K.; Wang, Y. Estimation of Origin and Destination Information from Bluetooth and Wi-Fi Sensing for Transit. Transp. Res. Rec. J. Transp. Res. Board 2016, 2595, 11–17.

[10] EMEL, G.G.; TAŞKIN, Ç.; AKAT, Ö. Profiling a Domestic Tourism Market by Means of Association Rule Mining. Anatolia 2007, 18, 334–342.

[11] Law, R.; Rong, J.; Vu, H.Q.; Li, G.; Lee, H.A. Identifying changes and trends in Hong Kong outbound tourism. Tour. Manag. 2011, 32, 1106–1114.

[12] Shoval, N.; Isaacson, M. Tracking tourists in the digital age. Ann. Tour. Res. 2007, 34, 141–159.

[13] Fukuda, D.; Kobayashi, H.; Nakanishi, W.; Suga, Y.; Sriroongvikrai, K.; Choocharukul, K. Estimation of Paratransit Passenger Boarding / Alighting Locations Using Wi-Fi based Monitoring: Results of Field Testing in Krabi City, Thailand. J. East. Asia Soc. Transp. Stud. 2017, 12, 2151–2169.

[14] Peters, M.; Piazolo, F.; Köster, L.; Promberger, K. The Deployment of Intelligent Local-Based Information Systems (ilbi): A Case Study of the European Football Championship 2008. J. Conv. Event Tour. 2010, 11, 18–41.

[15] Lee, D.G.; Ryu, K.S.; Bashir, M.; Bae, J.-W.; Ryu, K.H. Discovering Medical Knowledge using Association Rule Mining in Young Adults with Acute Myocardial Infarction. J. Med. Syst. 2013, 37, 9896.

[16] Versichele, M.; de Groote, L.; Claeys Bouuaert, M.; Neutens, T.; Moerman, I.; Van de Weghe, N. Pattern mining in tourist attraction visits through association rule learning on Bluetooth tracking data: A case study of Ghent, Belgium. Tour. Manag. 2014, 44, 67–81.

[17] Bermingham, L.; Lee, I. Spatio-temporal sequential pattern mining for tourism sciences. Procedia Comput. Sci. 2014, 29, 379–389.

[18] Shoval, N.; Ahas, R. The use of tracking technologies in tourism research: the first decade. Tour. Geogr. 2016, 18, 587–606.

[19] Musa, A.B.M.; Eriksson, J. Tracking unmodified smartphones using wi-fi monitors. In Proceedings of the Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems - SenSys ’12; 2012; pp. 281–294.

[20] Tsubota, T.; Yoshii, T. An Analysis of the Detection Probability of MAC Address from a Moving Bluetooth Device. Transp. Res. Procedia 2017, 21, 251–256.

[21] Danalet, A.; Tinguely, L.; Lapparent, M. de; Bierlaire, M. Location choice with longitudinal WiFi data. J. Choice Model. 2016, 18, 1–17.

[22] Potortì, F.; Crivello, A.; Girolami, M.; Barsocchi, P.; Traficante, E. Localising crowds through Wi-Fi probes. Ad Hoc Networks 2018, 75–76, 87–97.

[23] Abedi, N.; Bhaskar, A.; Chung, E. Tracking spatio-temporal movement of human in terms of space utilization using Media-Access-Control address data. Appl. Geogr. 2014, 51, 72–81.

[24] Kusakabe, T.; Yaginuma, H.; Fukuda, D. Estimation of bus passengers’ waiting time at a coach terminal with Wi-Fi MAC addresses. Transp. Res. Procedia 2018, 32, 62–68.

[25] Smallwood, C.B.; Beckley, L.E.; Moore, S.A. An analysis of visitor movement patterns using travel networks in a large marine park, north-western Australia. Tour. Manag. 2011, 33, 517–528.

[26] Wang, F.; Li, K.; Duić, N.; Mi, Z.; Hodge, B.M.; Shafie-khah, M.; Catalão, J.P.S. Association rule mining based quantitative analysis approach of household characteristics impacts on residential electricity consumption patterns. Energy Convers. Manag. 2018, 171, 839–854.

[27] Rong, J.; Vu, H.Q.; Law, R.; Li, G. A behavioral analysis of web sharers and browsers in Hong Kong using targeted association rule mining. Tour. Manag. 2012, 33, 731–740.

[28] Park, S.H.; Jang, S.Y.; Kim, H.; Lee, S.W. An association rule mining-based framework for understanding lifestyle risk behaviors. PLoS One 2014, 9.

[29] Fournier-Viger, P.; Chun-Wei Lin, J.; Kiran, R.U.; Koh, Y.S.; Thomas, R. A Survey of Sequential Pattern Mining. Data Sci. Pattern Recognit. 2017, 1, 54–77.

[30] Wei, Y.; Yang, R.; Liu, P. An improved Apriori algorithm for association rules of mining. In Proceedings of the 2009 IEEE International Symposium on IT in Medicine & Education; IEEE, 2009; pp. 942–946.

[31] Tank, D.M. Improved Apriori Algorithm for Mining Association Rules. Int. J. Inf. Technol. Comput. Sci. 2014, 6, 15–23.

[32] Mooney, C.H.; Roddick, J.F. Sequential pattern mining -- approaches and algorithms. ACM Comput. Surv. 2013, 45, 1–39.

[33] Pitchayadejanant, K.; Nakpathom, P. Data mining approach for arranging and clustering the agro-tourism activities in orchard. Kasetsart J. Soc. Sci. 2018, 39, 407–413.

[34] Hwang, Y.-H.; Gretzel, U.; Fesenmaier, D.R. Multicity trip patterns. Ann. Tour. Res. 2006, 33, 1057–1078.

[35] Buhalis, D. Marketing the competitive destination of the future. Tour. Manag. 2000, 21, 97–116.

[36] Raun, J.; Ahas, R.; Tiru, M. Measuring tourism destinations using mobile tracking data. Tour. Manag. 2016, 57, 202–212.

[37] Crouch, G.I.; Ritchie, J.R.B. Tourism, Competitiveness, and Societal Prosperity. J. Bus. Res. 1999, 44, 137–152.

[38] Franzoni, S.; Bonera, M. How DMO Can Measure the Experiences of a Large Territory. Sustainability 2019, 11, 492.

[39] Ryan, C. Tourism and marketing - a symbiotic relationship? Tour. Manag. 1991, 12, 101–111.

[40] Gunn, C.A. Tourism Planning: Basics, Concepts, Cases; 3rd ed.; Taylor & Francis, 1994; ISBN 0844817430, 9780844817439.

[41] Farrell, B.; Twining-Ward, L. Seven Steps Towards Sustainability: Tourism in the Context of New Knowledge. J. Sustain. Tour. 2005, 13, 109–122.

[42] Frechtling, D.C. Forecasting Tourism Demand: Methods and Strategies; 1st ed.; Butterworth-Heinemann: Great Britain, 2001; ISBN 9781136402425.

[43] Poon, A. Tourism, technology and competitive strategies.; CAB INTERNATIONAL: Wallingford, UK, 1993; ISBN 0851987516.

[44] Gopikaramanan, R.; Rameshkumar, T.; Senthil Kumaran, B.; Ilangovan, G. Novel control methodology for H-bridge cascaded multi level converter using predictive control methodology. Glob. J. Pure Appl. Math. 2015, 11, 3053–3071.

[45] Tan, P.-N.; Kumar, V.; Srivastava, J. Selecting the right objective measure for association analysis. Inf. Syst. 2004, 29, 293–313.

[46] North, M. Data Mining for the Masses; 2012; ISBN 9780615684376.

[47] Shou, Z.; Di, X. Similarity analysis of frequent sequential activity pattern mining. Transp. Res. Part C Emerg. Technol. 2018, 96, 122–143.

[48] García, E.; Romero, C.; Ventura, S.; Calders, T. Drawbacks and solutions of applying association rule mining in learning management systems. In Proceedings of the Proceedings of the International Workshop on Applying Data Mining in e-Learning (ADML 2007); Crete, Greece, 2007; pp. 13–22.

[49] Hahsler, M. arulesViz: Interactive Visualization of Association Rules with R. R J. 2017, 9, 163–175.

[50] Hahsler, M.; Chelluboina, S.; Hornik, K.; Buchta, C. The arules R-Package Ecosystem: Analyzing Interesting Patterns from Large Transaction Data Sets. J. Mach. Learn. Res. 2011, 12, 2021–2025.

[51] Du, Y.; Zhao, M.; Fan, G. Research on Application of Improved Association Rules Algorithm in Intelligent QA System. In Proceedings of the 2008 Second International Conference on Genetic and Evolutionary Computing; IEEE, 2008; pp. 414–417.

[52] Brin, S.; Motwani, R.; Silverstein, C. Beyond market baskets. ACM SIGMOD Rec.

1997, 26, 265–276.

Chapter 4

Assessment of The Impact of Snowfall on Travel Time Reliability Considering Different Road Types Using ETC2.0 Probe Data

4.1 Introduction

Travel time reliability (TTR) of transportation systems is an important issue uses to evaluate of successfulness of the transport operations. Informed road users can organize an alternative schedule due to unexpected events, such as national disasters, traffic congestions or adverse weather conditions with the highest possible accuracy. The reliability of travel time contributes to social mobility improvement, as the value of time on businesses and activities is increasingly important in the globalized era. Similar to the climate fluctuation, it is one of transport reliability influencing factors around the world.

In the past decade, the intelligent transport systems (ITS) plays a huge role in transport development. The technology provides an opportunity to understand the characteristics of traveler behaviors and traffic operations on road networks since the technology, electronic toll collection system (ETC), occupied the major high standard road network globally. The system was originally designed for road pricing strategies and congestion reduction in toll area during rush hour instead of the traditional toll fees collecting system. The toll data is being used in traffic operation study around the world. Likewise, the collective ETC2.0 probe data collected from intercommunication between authorized users and traffic operators has the brilliant capability to capture various travelling activities associated with a global positioning system (GPS). The ETC system provides a variety of information to road users (e.g. route selection, congestion avoidance, and disaster notification support) while collecting specific travel data (e.g. duration, longitude, latitude, vehicle ID, time stamps, and so on). Moreover, this technology is useful for transport utilization studies such as travel behavior investigation, traffic improvement and prediction, and also an estimation study for transport planning, corresponding with a generally significant variable on traffic studies such as speed, travel time, location in latitude and longitude coordination.

Regarding weather impact investigation on traffic operation, several past studies of weather impact addressed travel time, and speed is a significant element of a transportation system [1]. Various transport improvement studies involved the weather conditions topics such as rain and snowfall amount effect, and even temperature rising may affect mobility condition. Variation in travel behavior and transport systems performance in different seasons was partly explained by weather variation [2]. The adverse weather is a greater influence on travel pattern than the decline of roadway performance [3]. The weather conditions are an important factor as the congestion's root causes [4], as well as the TTR, one of the most important value in travel demand studies [5]. These explorations were reflecting the effects of weather conditions on traffic conditions, especially the variation of travel time but the studies hardly involved the hierarchy of road network.

Adverse weather conditions such as heavy rain, snow, ice, and fog can significantly affect travel patterns, road traffic operation, especially driving behavior and traffic safety. The extreme weather conditions can cause difficulty in normal driving. Particularly, the snow would result in a reduction in pavement resistance and vehicle stability; the icy road surface also makes a difference in controlling the vehicle. Moreover, the adverse weather may increase the intensity of traffic congestion, on the other hand, the driver will face the abnormal driving situations that mostly affects the ability to see and drive control, also has an impact on speed and travel time. According to the snowy conditions, roadway suffered when many road surfaces are most slippery because the moisture mixes with dust and snow powder that forms the icy road. The slippery roads can reduce traction and increase the risk of losing control of the vehicle because snow powder is not removed from the roadway. The travel time variability concern of traveler is greater than congestion of recurring on the roads network because it is typically unexpected [6]. The TTR is used to evaluate road network service quality or level of service (LOS), and it is reflecting the traffic mobility on the road network.

This paper discusses how to estimate travel time from the ETC2.0 probe data using comparison method throughout TTR viewpoint, regarding various scenarios of selected road hierarchy and weather impact integration, which several indices are used to examine the reliability performance with these data. The methodology is useful for evaluating travel time influenced by weather condition, especially snowfall conditions. The results show that it was possible to use toll data collections for travel time estimation and time reliability representation on road hierarchy including weather impacts, and it would help travelers to decide an appropriate schedule to avoid losing time on their daily journey.

4.2 Literature review

4.2.1 Reliability studies

The reliability measure in transportation has been reported in various studies over decades, by using time-based parameters of highway performance such as the standard deviation, travel time index, planning time index, and buffer time index, which generally proposed to deal with travel time reliability prediction [7]. These indicators often described in the viewpoint of travel time by considering a route segment of origin to destination marks [6]. A variety of traffic indicators used to determine the operation performance, but traveler mostly concerns about actuality travelling factors while thinking about travel plan (e.g. the delay, speed, travel time, and reliability) more than technical terms (e.g. traffic volume, capacity, density, and LOS). The road network which provided a high level of LOS also reflects the level of TTR. The external events also measure reliability as an influence of travel time when the most influential event is identified, a trip with under the average congestion level was showed [8]. TTR can explain in term of how travel time varies over the time considered, and a traveler likes to mention their experience in term of extra time on normal travel condition (e.g. it takes about 15 minutes than usual to the hospital this morning) [4].

However, researchers discuss how to clarify and measure TTR accurately. The TTR is an important indicator that helps the traveler to adjust their trips to the destination, and it very important for business travelers to avoid getting to meeting late, logistics distributor can maintain just-in-time service concept [9]. The TTR measurement is purposed to become a suitable indicator of congestion levels and reliability, and it used to evaluate the performance of widely road network. Moreover, time reliability can be defined as congestion levels [10]. It is a practical indicator of user congestion in transport experience at the consideration periods of time as well as the congestion level has been declared in several studies to explain the reliability of travel time. Hence, time reliability and congestions measurement are the important measurements of service quality for travelers because they do not consider only the average travel time but also the reliability of choosing route [8].

In addition, traffic congestion is simply defined as a longer time taken to get to a destination than usual, while congestion also reported increases in the amount of time added to reach activities. It is related to an excess of vehicles on a portion of the highway resulted in the decrease of speed and the rise in traffic volume, especially big city with high population and traffic density. The rising of travel time may involve the variations of the traffic congestions and the predictable variation as traffic congestion events or unpredictable as traffic accidents. The unpredictable variations connected to the unreliability of travel time travelers may miss predicting their exact travel time before departure caused by the uncertain environment [5].

Another framework to understand traffic congestion was purposed with a three-elements as a congestion results in a decrease of speeds and accessibility. It attracted the mass activities from nearby roadway and experiences. These elements will have an effect on choices about where to live, work, and play [11]. Additionally, the previous study see Figure 4.1 described the congestion as the result of seven root cause such as road capacity, traffic incidents, work zones, weather, traffic control devices, special events, and fluctuation in normal traffic [4]. [12] addressed viewpoints of unreliability in travel time considered; such as the variations of seasons and days of the week, travel changes due to weather conditions, and individual of traveler attributed, [5] argued that the travel time might be split into two components; free-flow time and additional time, which refers to an amount of time it takes a driver to arrive at the destination without facing any traffic and the latter refers to increase of travel time due to variations in the traffic conditions.

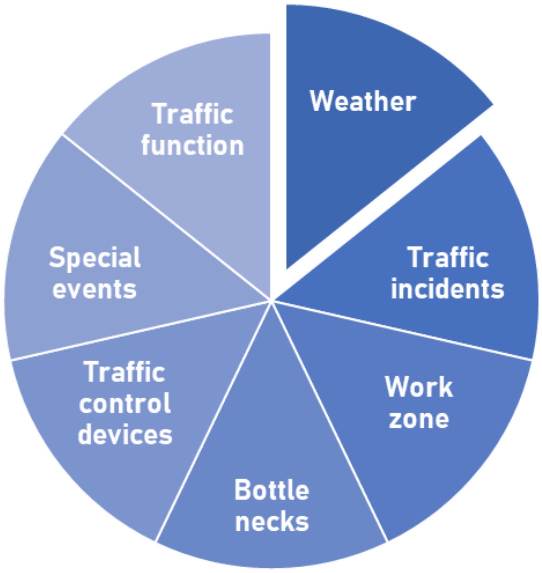


Figure 4.1 The seven root causes of congestion

4.2.2 Climate influences

The literature search was accomplishing to identify the article related to weather condition includes their effect on traffic operations. Since the climate effects on traffic operation become an interesting topic for decades, but a few studies were considering specific weather condition. The effect of weather conditions (e.g. rain, snow, ice, fog, and wind) on the TTR was measured in a span of travel time between 90th to 10th percentile, and the result had reported an increase of travel time variance. [2] also evaluated the LOS of a road network using macroscopic analysis method under rainy conditions [13]. The weather and environment along the roadway are essential parameters which also contributes to congestion and travel time variability [14].

The general perception of adverse weather as snowfall with traffic condition was worse and scary, and travelers may concern about trip planning (e.g. departure time when snow come, travel time span on the daily journey), because of climate fluctuation would impact travel time adjustment and traffic operation [15]. [3] The reports on weather and traffic condition study reveal the impact of light snow results in an 8% increase of travel time and 10% during heavy snowfalls. Heavy snowfall event has a potentially huge impact on the operating of speed thus travel time would also be affected and estimated declines in free flow speeds [16]. Likewise, the report on adverse weather such as heavy snow and rain have significant negative impacts on traffic operations [13]. In addition, the impact of rainfall on freeway affects traffic flow characteristics. When compared with normal weather, it resulted in a change under adverse weather [17]. [18] argued that snow depth is one of the most influential factors on winter weather condition study, and it is used to evaluate the reliability of travel time during the winter season as well.

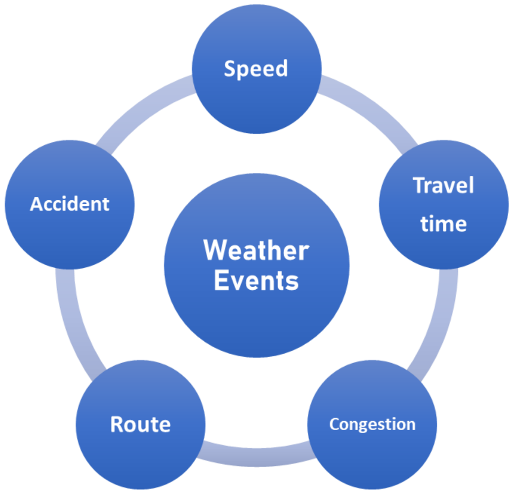


Figure 4.2 Weather events cause of congestion and unreliable travel

In the high standard road network, traffic operators apply ITS system to inform unclear environment as congestion involved weather conditions, which is the warning system that helps travelers choosing under a certain environment for their trip planning similar to [19] the proposed adaptive strategies and algorithms that integrate weather impact with toll data to forecast travel time. The perception of weather event affected on travel time is essential to make precise estimations and provide additional information to traffic operators [2]. The effect of weather conditions on highway can be categorized by considering the traffic operation parameters (e.g. volume, speed, and headway), and studies which consider the impact of weather events (e.g. rain, snow, fog, smoke, and extreme temperatures) on the quality of traffic services (e.g. delay, level of service, congestion rate, reliability) [18]. However, the weather variability effect on travel time has been reported to have slightly scaled, especially when compared with the number of studies dealing with the freezing traffic condition in snowfall events. The changes in travel time in relation with snowfall concerning the road hierarchy also have not been established well

4.3 Data collection and extraction

4.3.1 ETC2.0 probe data

The ETC system is an ITS technology for the toll roads, and its original purpose is to avoid congestions and to service disruptions at toll gate [20]. In the past decades, the new generation of toll collection system "ETC2.0" has been proposed together with an advanced facility system by ITS Technology Enhancement Association (ITS-TEA) Japan's traffic administrator see Figure 4.3. The ETC2.0 system provided advantage information to allow road users for trip planning procedures such as travel time estimation, road network congestion observation, and real-time travel information sourcing. As aforementioned, technology has made possible to estimate variance in travel times on road network using a data collector device. However, achieving an accurate estimate is difficult because the ITS device as a detector only to monitor traffic condition at a single point as it may present different results of the whole road section [19]. There is a high possibility to obtain traffic trajectory and travel time data along the road network since the cars are equipped the ETC devices.

This study used electronic toll collection data, ETC2.0 probe data, to evaluate travel time and reliability trends over consideration periods in association with the road hierarchy classification system and historical weather data in the Sapporo city area. The toll data obtained individual historical travel data from road network, including specific information such as position in latitude and longitude coordination, speed, car ID, travel time, travel distance, and so on. Collective datasets have been mapped on the digital road map database (DRM) using a map matching technique. ETC2.0 probe data divided into five hierarchy of road network that illustrated in Figure 4.4 as well as the Sapporo road map database shown in Figure 4.5. It consists of digital cartographic data in which location and various information, obtained hierarchy layers of the road network and vehicle probe data. This study targets on the vehicles that use road network including service area more than few minutes, the data around trip starting and ending at the same point have deleted. Map matching procedure has applied to the road network map, associated with road hierarchy categorization, which represents the layer of road hierarchy individually see in Figure 4.6

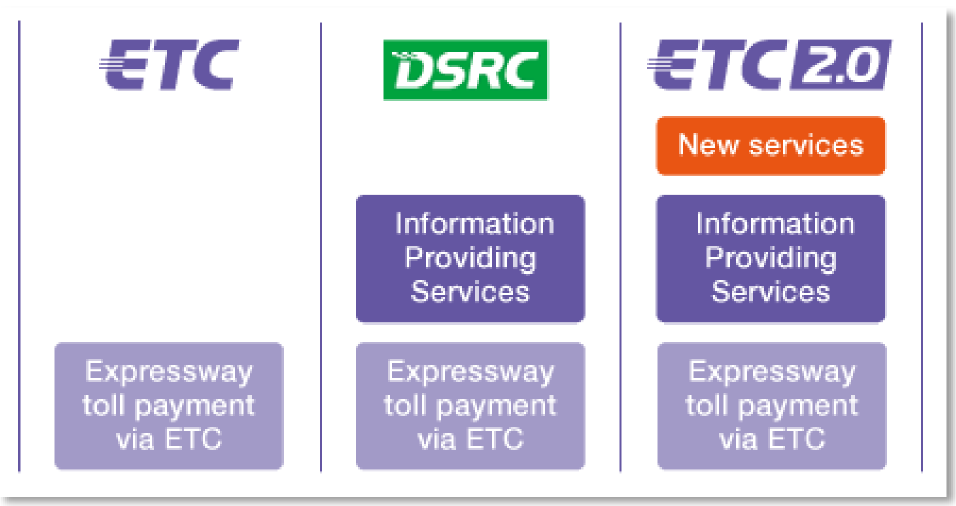


Figure 4.3 ETC2.0 as a modern intelligent transport system

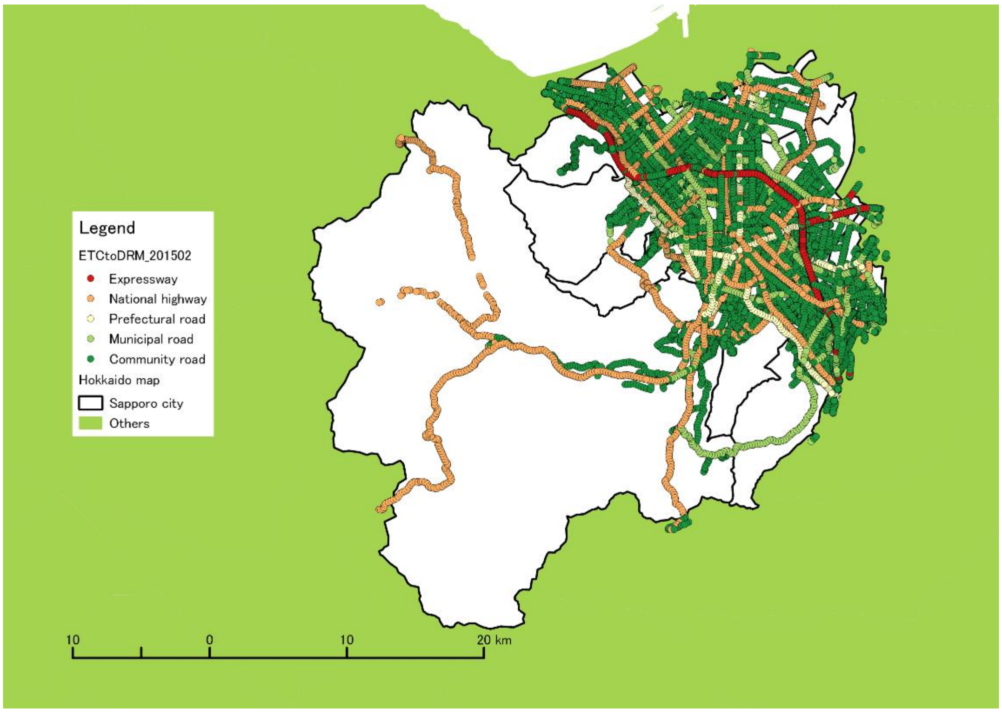


Figure 4.4 ETC2.0 probe data illustrating road hierarchy

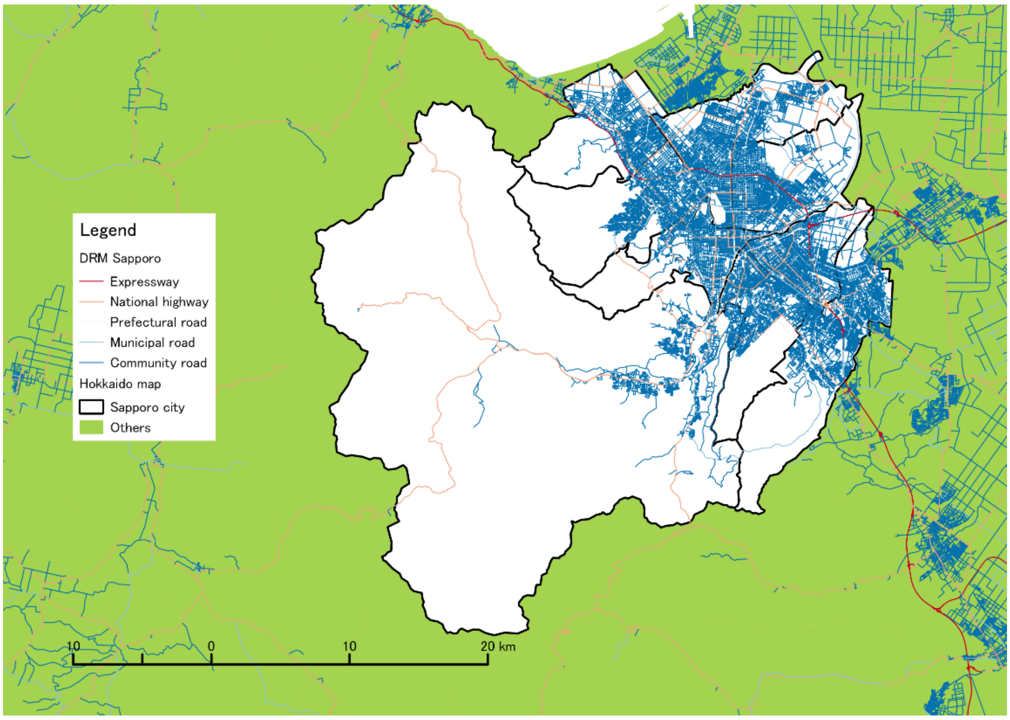


Figure 4.5 Sapporo DRM categorized by road hierarchy

Expressway

National highway

Prefectural road

Municipal road

Community road

Figure 4.6 ETC2.0 probe data represents the road hierarchy layers

4.3.2 Road hierarchy

To evaluate the travel time on road stratum of the entire city of Sapporo, all the link in the city based on Hokkaido's road hierarchy system presented by Hokkaido Regional Development Bureau under Ministry of Land, Infrastructure, Transport and Tourism (MLIT) [21]. Road hierarchy is divided into five categories (e.g. expressway, national highway, prefectural road, municipal road, and community road), which are related to access management as it provides access to land development and involves with mobility and accessibility concepts. The road network coverage and amount of effective data have been shown as sample size (number of trips used for analysis) separated by seasons; a sample size of winter showed smaller amount than non-winter seasons regarding the original data limitation as shown in Table 4.1.

The classification of road hierarchy is represented into five different layers as mentioned, it is represented by relation of mobility and accessibility principal see in Figure 4.7. The highest hierarchy, the expressway is a high standard highway connecting major cities and transport facilities which is exclusively designed for high speeds and full access control, equipped with road pricing system. The national highway network is acting like a backbone of the city, and it expects to carry out a large scale of traffic volume through the city without charge. Meanwhile, the other road functions have a greater capability of access than a high level of road hierarchy but less mobility. The lower hierarchy such as prefectural, municipal, and community roads respectively contribute traffic from household to high mobility road level in the morning and backward in the evening generally.

Table 4.1 Sapporo road network ETC2.0 configuration and characteristics.

Road hierarchy	H1	H2	H3	H4	H5
Description	Expressway	National highway	Prefectural road	Municipal road	Community road
Length, km.	38.00	152.26	239.82	56.94	5,223.86
Data coverage, %	100	75.54	47.93	100	42.38
Sample size, (number of trips)					
Winter	804,612	1,901,187	926,714	1,337,132	3,463,088
Non-winter	1,565,014	3,337,141	1,722,548	2,475,415	6,198,934

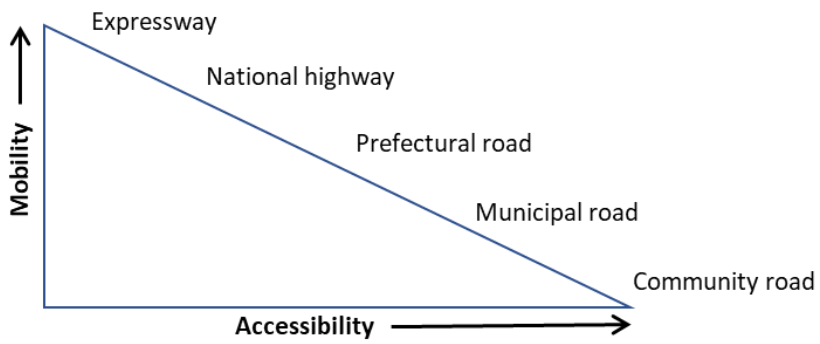


Figure 4.7 Mobility versus accessibility by hierarchy of the roads

4.3.3 Meteorological data

This paper used the daily weather data collection provided by the Japan Meteorological Agency (JMA) [22]. According to the purposes of this study, to declare the snowfall effects on the travel time basis, the integration of weather data to electronic toll collection probe data is necessary. Comparison analysis between winter and no-winter is a majority to examine the phenomena of TTR evaluations on road hierarchy. Hence, the weather data including a branch of climate perception such as snow depth, temperature, wind speed, precipitation, and humidity are on a daily average measure basis. In this research, only the snow depth record in centimeter per day is used to achieve the purpose of the study. The snow data collected from Sapporo climate observatory station (WMO Station ID:47412 Lat 43.036' N Lon 141.197' E) reports daily average snowfall coverage in the Sapporo's metropolitan. It provided snowfall records in Sapporo city since 1989 shown in Figure 4.8. Therefore, the consideration period during the winter season with snowfall record and non-winter seasons without snowfall, while the target months was selected from the years 2015-2016 regard available data. The selected months in both periods of winter and nonwinter shown in Table 4.2, there were obtained the comparison analysis further. In addition, the record also representing snow depth, which condensed the average daily snow over months. Amount of snow depth was trending downwards, and random appearance see Figure 4.9.

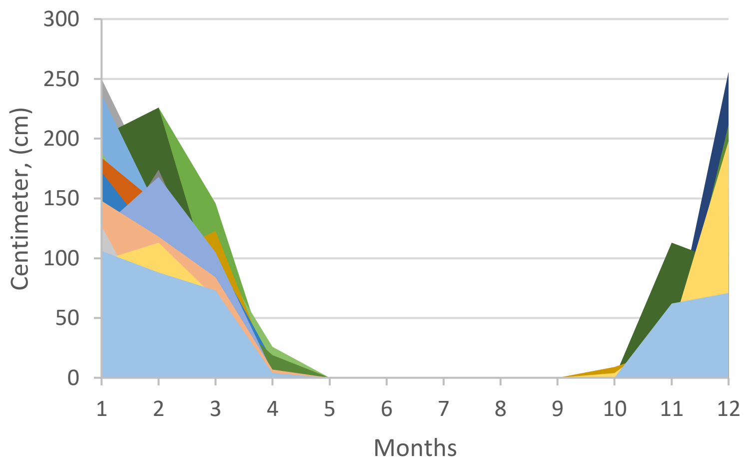


Figure 4.8 Snowfall record of Sapporo city along the year, period of 1989-2017

Table 4.2 Selection of months.

Seasons	Months			
Winter	December	January	February	March
Non-winter	June	July	August	October

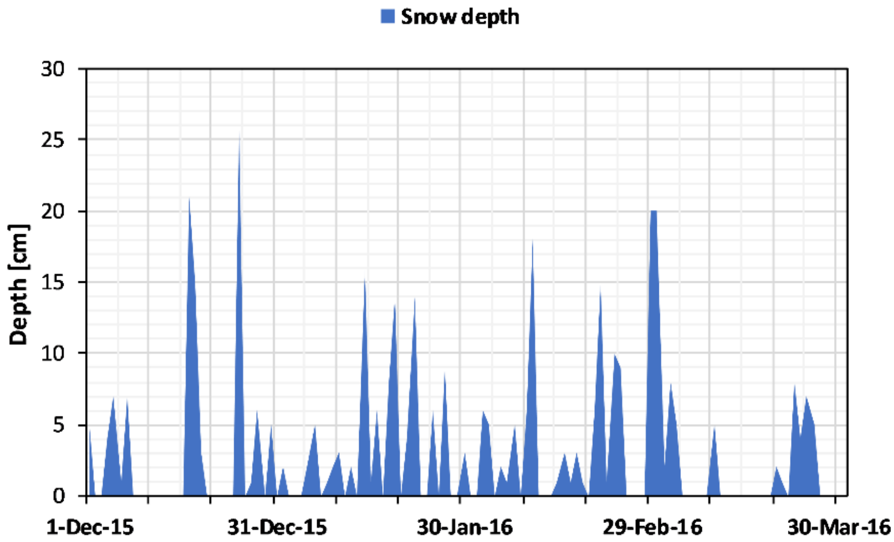


Figure 4.9 Daily snow depth within selected study period

4.3.4 Evaluation of time reliability

In the highway performance theme, previous researches have discussed various reliability measurements [5] [6] [23]. The investigation of weather condition impacts on the traffic operations, traffic flow characteristics (e.g. free-flow speed and average travel times) are considered, and the target area is focusing on the climate effect on a simple living road network. Furthermore, time reliability evaluation was considered as a roadway LOS performance indicator [18] [24] [25]. In this study, the data analytics and statistical methodology have occupied electronic toll collection data, the performance indicators [4] [26] such as standard deviation, average travel time, planning time index, and buffer time index have performed for TTR evaluation on road hierarchy procedure [7]. The TTR measurement was frequently described in the previous researches in term of travel time (minutes) for roadway segment of a given length (kilometers). The evaluation analysis acquired a general knowledge of travel time required besides consideration periods, and a map matching technic was applied to treat ETC2.0 probe data as it was matching sequential travel patterns and was gathering real-world coordinate into a digital map. Thus, it has classified GPS log data into highway categories under specific conditions; divided trips data by road hierarchy classification, checked the required time distribution on each stratum and performed the seasonal comparisons; observation in hourly units and clarified the time zone during the day, which time reliability decreased from yearly seasonal comparison to snowfall period. Next, to analyses the influence of the snow scale in details, the period is classified by the amount of snow, and the reliability index of each duration was obtained.

Standard deviation value is widely used in traffic studies to measure confidence in the statistical term to represent the variability of average travel times despite the general dataset. Because travel time on roadway segment varies widely, the standard division obtains a high value which represents poor reliability. On the other hand, the high level of reliability is recognized when most of the cars on the roadway segment take the same time [6]. Reliability indices in this study are stated in term of travel time basis.

The congestion index is a measure of congestion level in Equation 4.1. When the comparison between the average travel time conditions to FT is free-flowed, the travel time conditions $(\overline{TT})$ is the average of travel time in the specification period (in minutes per kilometer). The congestion index can be computed as

$$C_i = \overline{TT}/FT \tag{4.1}$$

In the planning time, PT, is simply a total time needed to be punctually arrived at 95 percent of the travel time, which represents how much total time travelers should plan to ensure on-time arrival (in minutes per kilometer) in Equation 4.2. The planning time index, PTi, shows how much larger the planning time is than the free-flow travel time as the ratio of 95 percent of time to travel in free flow condition. The measurement of reliability is amplified as

$$PT_i = PT/FT \tag{4.2}$$

The buffer time, BT, represents the extra time travelers should add to their average travel time to ensure an on-time arrival determine by planning time minus average of travel time (in minutes per kilometer) in Equation 4.3. Buffer time index can be represented as

$$BT_i = (PT - \overline{TT})/\overline{TT} \tag{4.3}$$

It interprets the amount of additional time needed to be on-time at 95 percent of the travel time. Indexing the measurement provides a time and distance neutral measurement, but the actual minute values could be used by an individual traveler for a trip length.

4.4 Results

This section discussed the influence of weather conditions on the travel time on a collection of the road hierarchy systems in Sapporo metropolitan as mentioned in the above in the methodology data section which discussed an extreme snowfall occurrence area. Hence, this study will evaluate the time reliability by comparing analysis of travel time distributions, time zone analysis, and snow depth effects as follows.

4.4.1 Comparison of reliability indexes

Table 4.3 shows time reliability indices which are classified by road hierarchy and seasons. The seasonal comparison analysis has seen greater indices of time reliability evaluation during winter than on non-winter selection. The winter's congestion index represented poor traffic congestions level and had gotten worse from a high level to a low level of road hierarchy respectively. The expressway obtained the lowest congestion level in both seasons when compared with community road as the lowest hierarchy. The community road has approximately two times larger congestion than the expressway during the winter period. They are showing a high possibility to experience traffic jam on low level rather than the high level of road hierarchy. Planning time indexing during winter shows the planning time is larger than the free-flow travel time in non-winter seasons, especially the community road index which presented the largest number and reflected travel on the lowest road hierarchies which may consume more time than other upper road hierarchies, and the travelling speed was much slower. The buffer time index has the most consistency among the high standard highway and municipal road, which is twice over the average travel time but increased on the lowest hierarchy. The huge effect on travel time occurred on the community road, and when comparing to the season, the winter index was found to be larger than non-winter.

Table 4.3 Time reliability measurement indexes.

Road hierarchy	Season	Std.Dev.	Ci	PTi	BTi
H1	Winter	1.668	1.61	5.41	2.37
	Non-winter	1.625	1.54	4.96	2.22
H2	Winter	3.174	2.02	6.77	2.36
	Non-winter	2.914	1.98	6.65	2.36
H3	Winter	3.329	2.11	7.84	2.71
	Non-winter	3.011	2.07	7.22	2.49
H4	Winter	3.320	2.59	9.83	2.80
	Non-winter	3.033	2.45	8.71	2.55
H5	Winter	4.487	3.26	16.33	4.01
	Non-winter	4.011	2.98	15.00	4.04

This represented the additional time needed in the winter over in the non-winter. Furthermore, the standard deviation indicated that the high standard highway has a smaller value than ordinary road hierarchy as it represented a stable hierarchy with small variations.

Moreover, the impact of snowfall has been captured as such a great value during the winter season. Time required comparisons between seasonal divided by road hierarchy have shown the extra time needed to be on-time in both seasons shown in Figures 4.10 and 4.11. The highest road hierarchy such as the expressway reports a less than two minutes buffer time needed per kilometer in both seasons, and the extra time increased respectively on the ordinary road hierarchies as well as the planning time on the expressway was smaller than others. Prefecture and community roads have seen 21% and 18% rise in both buffer time and planning time respectively, and when compared with other road hierarchies, it was greater than the expressway which reported a 15% increase on the buffer time as well as an increase of 12% in the planning time when reaching the winter. The national highway has a small change of less than 10% increase during winter when compared to non-winter in both of time-required indicators. Regarding the largest extra time on the system, the community road was found to have the largest spanning of buffer time and planning time, which was the lowest road hierarchy and also the smallest road functional in an urban area. Thus, these arguments represented the possibility to be arriving on-time to the destination if travel on the expressway because the time required indicators are smaller than others. It reflected appropriated time reliability and confirmed that amount of time required was increasing during the winter and approached the bottom of hierarchy systems.

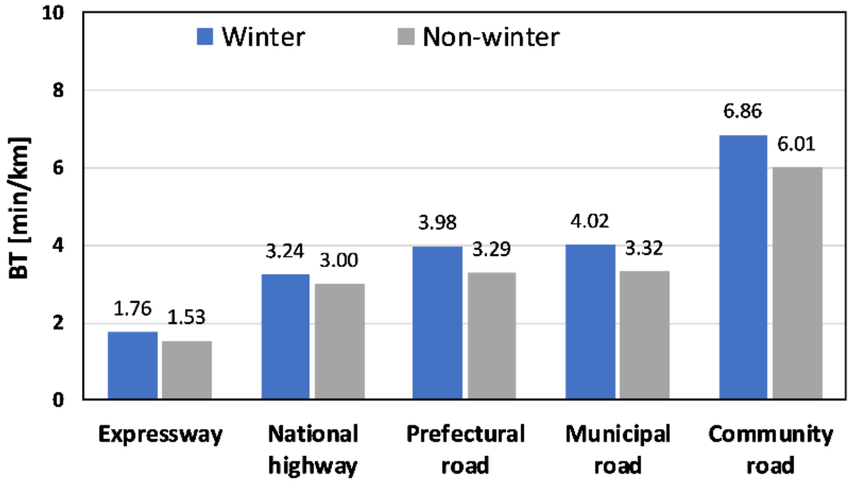


Figure 4.10 Buffer time comparison of seasons represent in minute of time

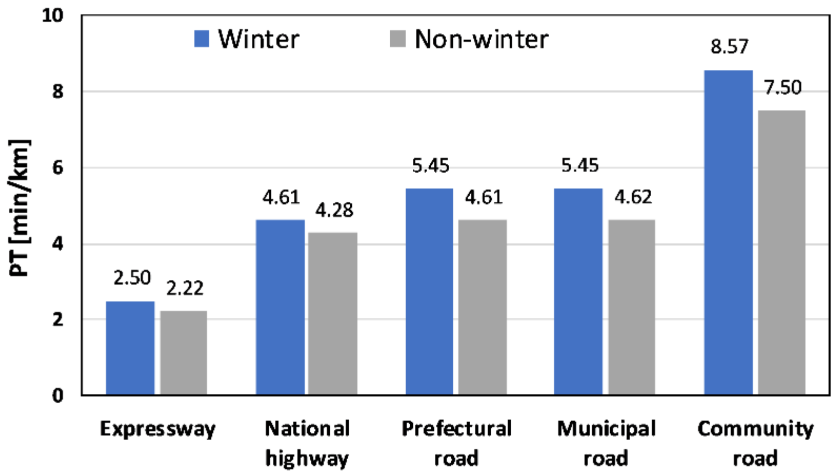
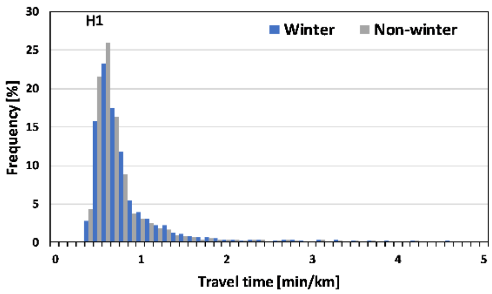


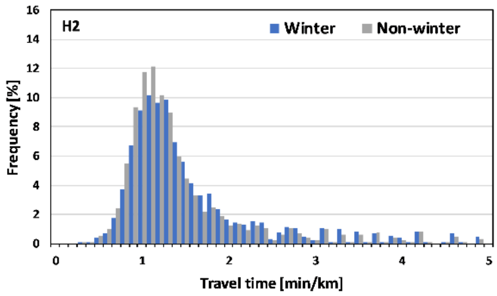
Figure 4.11 Planning time comparison of seasons represent in minute of time

4.4.2 Comparison of travel time distribution

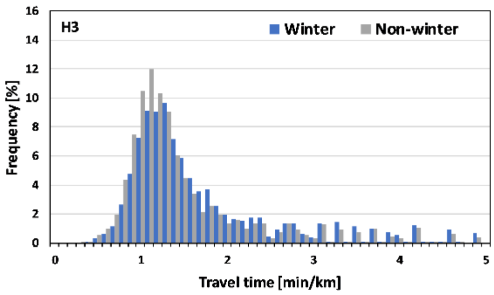
The distribution of travel time illustrates how travel time looks like on each road hierarchy during winter and non-winter seasons by using ETC2.0 probe data. Travel time attempted by travel distance and speed, and several inappropriate data sets were terminated as empty row (non-available data), duplicated data, and overlapped link. Multiple layers of road hierarchy are associated with trip distributions through each hierarchy layer, and the travel time was calculated over each road hierarchy sections to compare between winter and non-winter periods. Hence, the comparison analysis showed the divergent travel time needed by individual hierarchy along the season. The histograms have projected the distribution of travel time on the roadway orders when integrated with the weather conditions, it also exploded the time reliability along with velocity perspective shown in Figure 4.12.



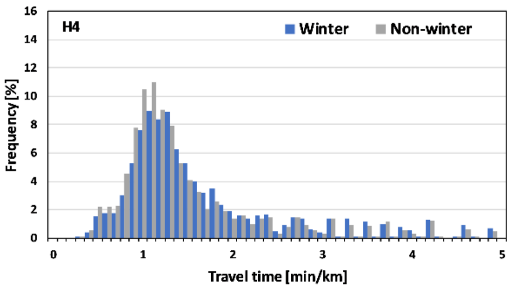
(a)



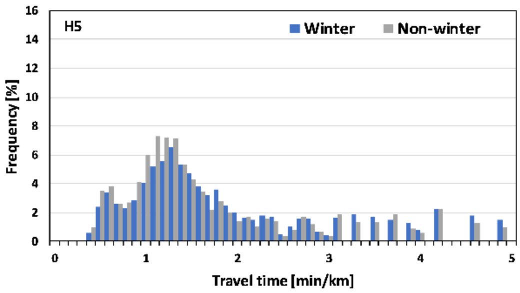
(b)



(c)



(d)



(e)

Figure 4.12 Travel time distribution with seasons comparison: *(a)*expressway, *(b)* national highway, *(c)* preflectural road, *(d)* municipal road, and *(e)* community road

According to the distribution of travel time per distance at 1 min/km, the travel time distribution of high standard highway such as the expressway can be seen as smaller than other hierarchies. The maximum distribution obtained was 0.6 min/km (100 km/hr), and as well as the narrow distribution's skirt, it indicated the high speed on the top hierarchy shown in Figure 4.12 (a). On the contrary of the lower road hierarchy, a gap was seen between winter and non-winter as it was the evidence of weather conditions influencing the roadway. Regarding the level of mobility and accessibility characteristics, the road hierarchy function is greatly different by order, for example, the expressway has full access control for speed travel. In the case of 2 min/km or more, the spread of time distribution trend increases right after 1 min/km even the distribution's skirt on the right side becomes larger. Likewise, in the 3 min/km case, the bunch of skewed positive data has ended, and the distribution has gotten smaller when the tail spreads to the right side with an outlier, while great winter effects were seen shown in Figures 4.12(b), 4.12(c), 4.12(d), and 4.12(e). Additionally, the distribution of travel time illustrated an evidence of the small amount of time needed on the expressway, and it also reflected the low travel speed on lower hierarchies as wide-skirt distributions. The results clarified the low hierarchy affected by winter condition as snow covered.

4.4.3 Comparison of time zone analysis

Above, the results point out the low road hierarchies affected from the influence of snow cover in the winter season. The comparison analysis was applied to clarify more details of time reliability during daytime targeting the time zone of the day to examine time reliability behavior. This section decided to cover the daytime from 6:00 to 18:00 hours, because of data limitation which obtained a small amount of the late night from the late afternoon data. It was unsatisfied for analyses. In each time zone (hourly), in order to see how much time reliability declines during snowy season are comparing with non-snow seasons, the ratio (growth rate) was determined by dividing the value of the snowy season by the non-snow seasons. Thus, the time required buffer time and planning time are illustrating in the bar graph.

As for the growth rate of the buffer time and planning time shown in Figure 4.13, it represented the buffer time growth ratio based on the winter season. The low hierarchies have seen a large ratio during the morning rush hour (6:00-8:00 hours). The community road and municipal road are especially affected at 6:00 hours by having obtained 1.27 and 1.23 times respectively compared to the winter while most of the road hierarchy suffered at 7:00 hours. Along the day period, the municipal road maintained a high ratio as well as community road. Surprisingly, the expressway suffered during the early afternoon of 1.27 times, which was greater than non-winter at 13:00 hour and decreased in the late afternoon. According to evening rush hour (16:00-18:00 hours), the high ratio was seen on the national highway and prefectural road regarding leaving time.

Likewise, Figure 4.14 represents the planning time growth ratio which generated a high level during morning rush hour and increased greatly at the time of commuting to work. The 6:00 hours obtained 1.25 greater times than non-winter on the lowest hierarchy such as the community road and municipal road which experienced the same trend. The group of low hierarchy level maintained a high ratio during the daytime. The expressway’s ratio was spiking in the early afternoon and was decreasing respectively into the late afternoon. The evening rush hour showed an increase of ratio on the national highway and the prefectural road which obtained about 1.18 times maximum in relation with the time of leaving in the evening. These explorations concluding the impact of snowfall on the low hierarchy has significant recognition.

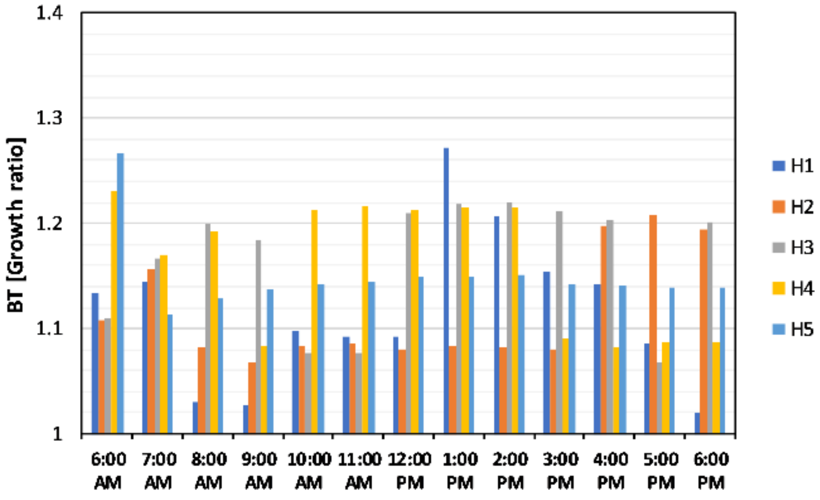


Figure 4.13 Buffer time growth ratio over daily time periods regrading non-winter based by road hierarchy.

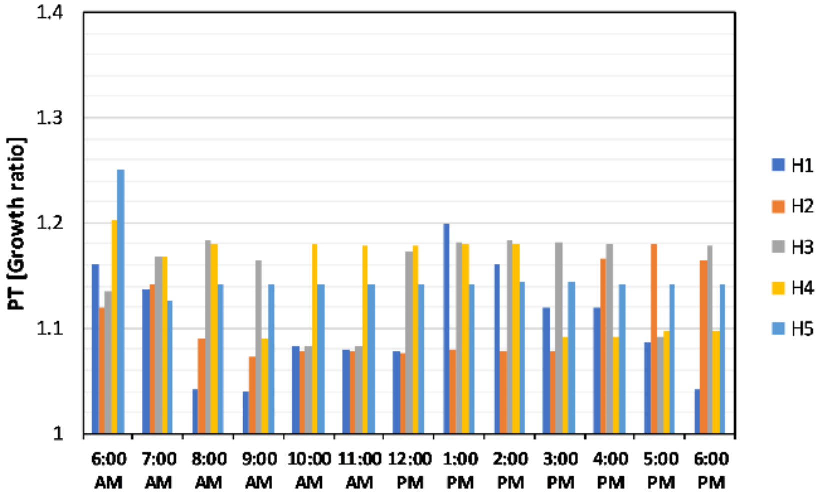


Figure 4.14 Planning time growth ratio over daily time periods regrading non-winter based by road hierarchy.

On the other hand, on the highway, it is hardly affected by snow coverage. It can be concluded that the daily time zone is susceptible to the influence of snowfall because the growth ratio is increasing at the time of leave. Thus, the extra time preparation is needed while travel during winter, especially on the low hierarchy of roadways to avoid late arriving at the destination due to the impact of snowfall, particularly during the rush hours.

In addition, the extra time illustration is representing the amount of buffer time shown in Figure 4.15 and planning time shown in Figure 4.16 on daily time periods basis. It captured the larger extra time during the winter season (DWIN) than non-winter season (NWIN), which also covered the entire road hierarchy system. The expressway has seen the smallest amount of extra time needed to be on time as it obtained about 2 min/km in maximum during the winter season. However, the lower hierarchy such as the national highway, prefectural road, and municipal road have performed a sustained level of the extra time of about 4-6 min/km maximum. These road hierarchies needed small extra time on the early morning (6:00 hours) when compared to others time zone. However, the huge gap occurred on the community road, and it suffered the huge extra time at about 6.7-8.6 min/km maximum, which was also greater than any others hierarchy. The rush hour was recognized as a peak period of traffic demand; when the general time to leave in the morning and afternoon, congestion concerns and extra time needed in preparing to arrive on time were considered. According to BT and PT, the morning (6:00-9:00 hours) and the afternoon (16:00-18:00 hours) demonstrated the rush hours period of seasons shown in Figure 4.17 and Figure 4.18.

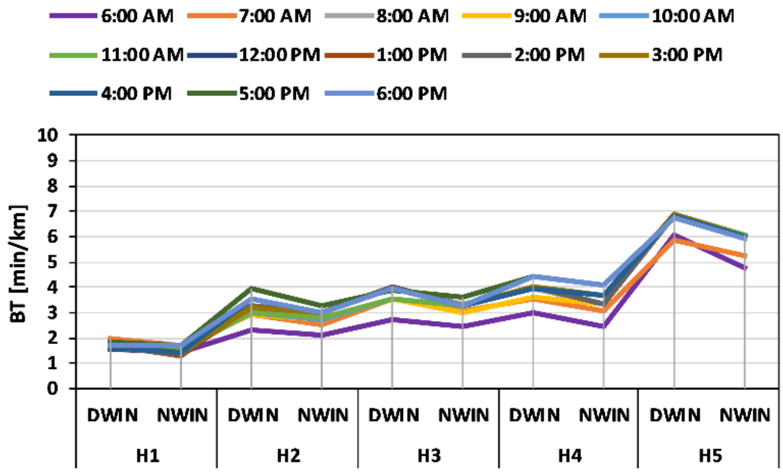


Figure 4.15 Buffer time due to seasons comparison and road hierarchy.

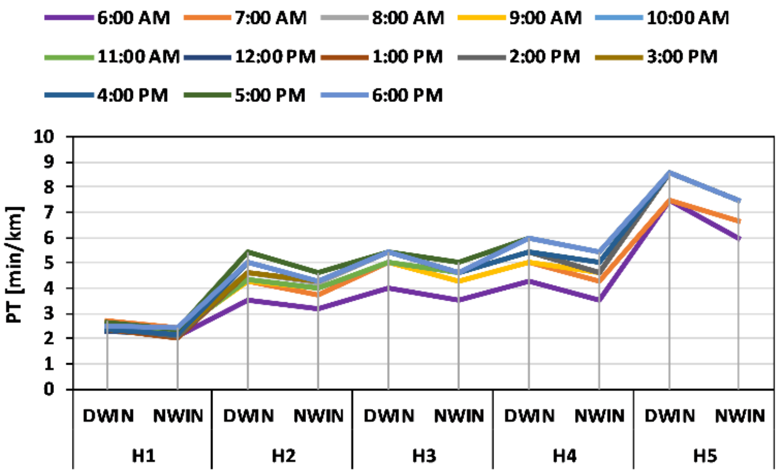


Figure 4.16 Planning time due to seasons comparison and road hierarchy.

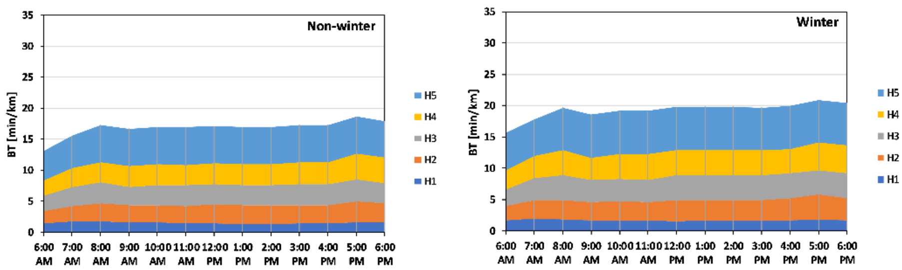


Figure 4.17 Buffer time by road hierarchy among daily time periods: *(a)* non-winter, and *(b)* winter.

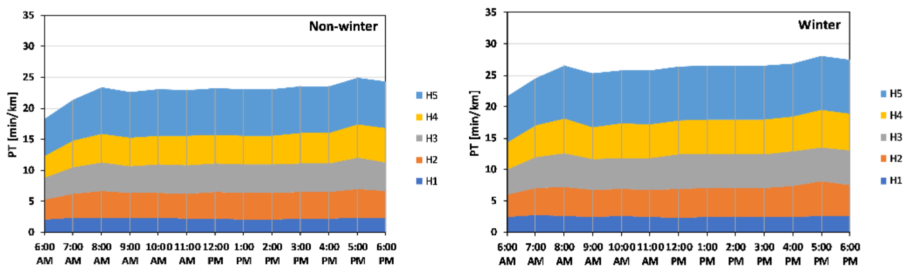


Figure 4.18 Planning time road hierarchy among daily time periods: *(a)* non-winter, and *(b)* winter.

Next, the null hypothesis test was performed to evaluate that travel time for daily time periods were statistically independent. Table 4.4 represents the probability of buffer time and planning time in seasons which suggests rejecting the null hypothesis confirming that there is no statistical dependence between a pair of travel time.

Table 4.4 The probability of hypothesis test for independence of finding correspondingly to daily time periods analysis (6:00 to 18:00 hours) between winter and non-winter season.

Road hierarchy	BT		PT	
	t-statics	*p*-value	*t*-statics	*p*-value
H1	6.402	1.697 E-5	8.140	1.574 E-6
H2	7.160	5.739 E-6	8.680	8.080 E-7
H3	9.630	2.626 E-7	13.120	8.894 E-9
H4	10.515	1.038 E-7	14.294	3.366 E-9
H5	21.078	3.776 E-11	28.151	1.250 E-12

The hypothesis test performed included the daily time periods and was classified into 13 hours from early morning through late afternoon as mention above. Thus, the degrees of freedom are 12. One-tailed test values were generated, and it was larger than the critical value which therefore rejected the null hypothesis. In addition, their significance probability values on an individual hierarchy were smaller than a significance level alpha of 5%. From these perspectives, the statistical probability results explained that the travel time among time periods of the day are statistically independent. In this section, the majority determined the snowy influence over daily time periods and the growth rate ratio of buffer time and planning time-based on the non-winter season provided — also, perspective perception on illustrating to the clarified amount of extra time needed over time, which suffered through the winter condition. The statistical process applied to assess dataset has practical significance.

4.4.4 Comparison of snow depth effects

This section explored the effect of snow depth on travel time on the road hierarchy. The experience has shown that the required time and its variation become large due to snowfall, and in the investigation of past research, there are a few studies that investigate on what degree the snowfall intensity can affect the roadway especially the order of road hierarchy.

In the period of winter season as the months' collection shown as Table 4.2., there were no snow every day during winter hence the snow depth record has seen missing data or zero snowfall on that day. Figure 4.19 represents the relation of daily snow depth record and buffer time which illustrates the degree of snow's effect on the road hierarchy realistically due to the snow record. The buffer time increased aligning with snow event occurrence properly, which greatly affected the low hierarchy. For example; due to the huge snow record of 26 cm at the end of December, the buffer time was skipping, particularly the ordinary roadways, and the high standard highway also increased at about 0.4 times from regular day without snow. Similarly, the planning time illustration demonstrates the relation of snow events and the total time that traveler should plan to ensure on-time as planning time shown in Figure 4.20. The planning time performed the high rising amount of time since the occurrence of snow events. These explorations explained that snow events affected on travel time, hence reflecting the quality of time reliability on the road hierarchy.

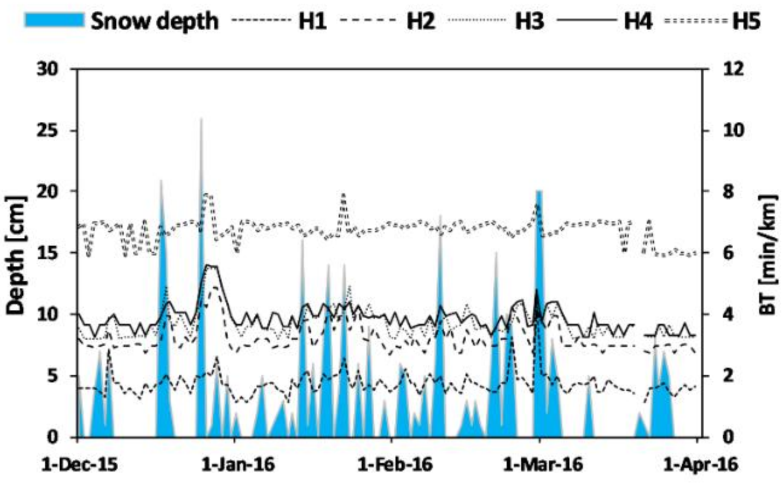


Figure 4.19 Visualize of snow depth and buffer time relation.

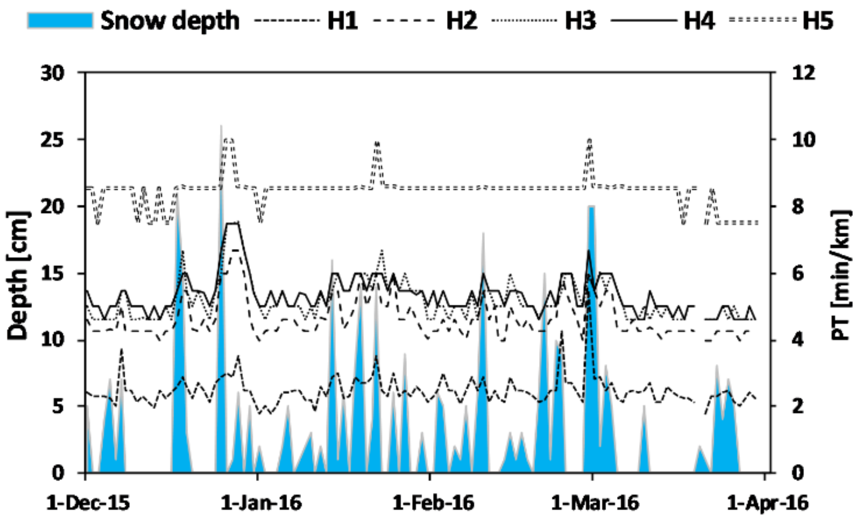


Figure 4.20 Visualize of snow depth and planning time relation.

To determine the impact of snowfall on the road hierarchy even when the period is divided by daily snowfall amount, the trips were extracted and totaled according to the road hierarchy. The standardize values have been generated by having obtained 0 cm snowfall amount as 1 ratio which represents the relation between BT and PT with snow depth shown in Figure 4.21 and 4.22 respectively. The snow effect tends to increase and align with the increase of snowfall; as for a small amount of snow as 1-2 cm resulted in a 0.98 times ratio on the expressway. It was less affected when compared to others road hierarchy at the same level of snowfall, and the impact has increased to 7 cm at about 1.2 times ratio which later declined. At about 10 cm, there were small effects on prefectural road and community road at about 0.97 times ratio. The great effect occurred at 14 and 26 cm on the expressway and national highway. As for 20 cm, it has fewer snow effects on all hierarchy, and the national highway was affected the most at 26 cm. In summary, the most effects occurred on the national highway also on a municipal road. Although the fluctuation of snowfall is an uncontrollable factor, the exploration clarified the existing of snow influence on TTR by road hierarchy.

Moreover, linear regression trend line to visualize the overall effect of snow depth due to a specific period on BT and PT have been illustrated. The trend of the last hierarchy shows the smallest slope of the least-square regression line when compared to others hierarchy, and it hardly affected while snow events occurred. On the other hand, the national highway has the largest slope of the least-square regression line, and it was representing great impact while snow events occurred; corresponded with the above result of the greater effect on the national highway.

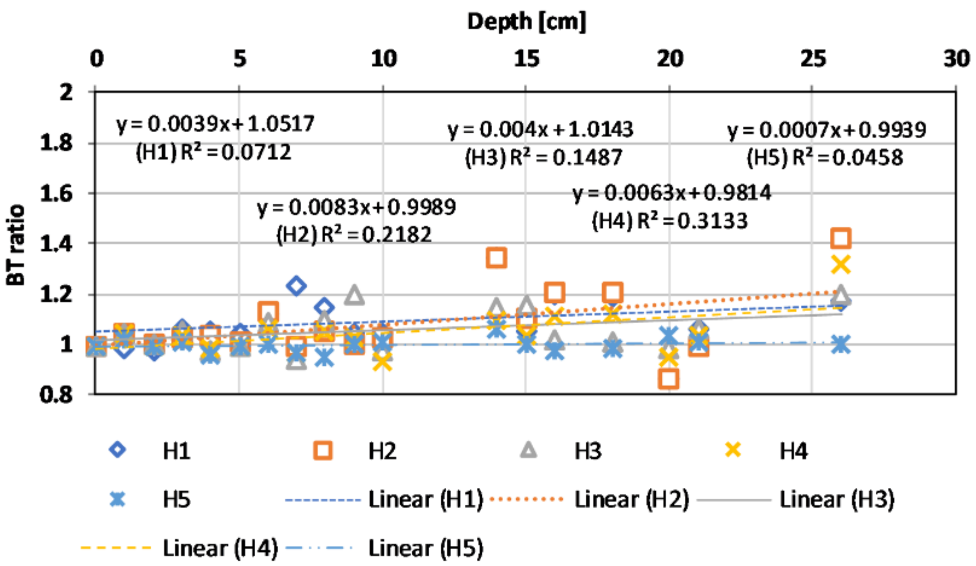


Figure 4.21 The relation of snow depth and buffer time ratio

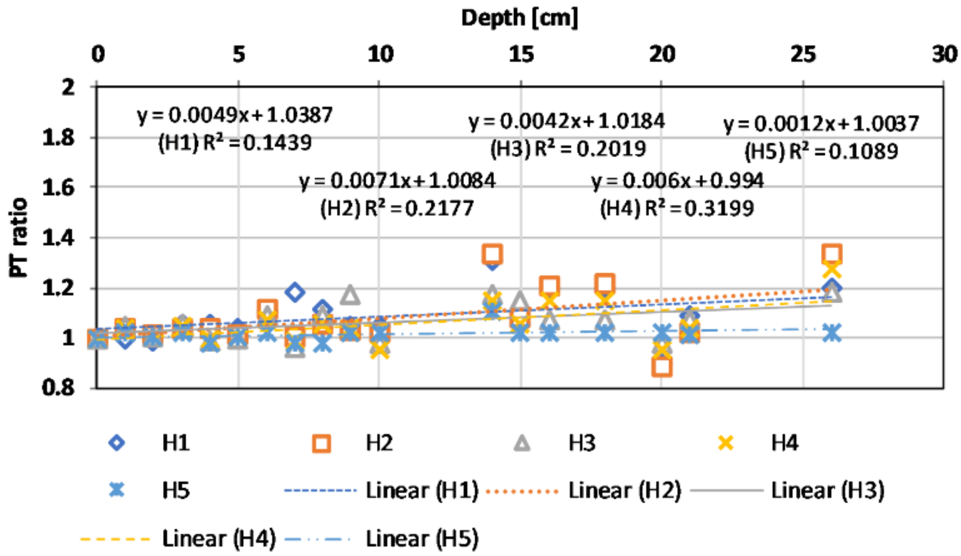


Figure 4.22 The relation of snow depth and planning time ratio.

Furthermore, the analysis of variance applied to determine whether the five groups of the hierarchy are different from each other or statistically independent. The ANOVA test was mainly used to analyses all 18 observations of snow depth record in regard of matching procedures of average time required by road hierarchy together with snow depth in centimeter over a consideration period which captured snow depth at the minimum of zero cm and 26 cm maximum. Thus, hypothesis test was performed, and the degrees of freedom is (5-1) = 4 on both buffer time and planning time. Their significant probability values are smaller than a significance level of 5%, which the results have failed to accept the null hypothesis. It explained no statistic dependence of time required data between the group of road hierarchy shown in Table 4.5.

Table 4.5 ANOVA test for determine statistically independent related to snow dept record.

Source of variation	*Count*	*df*	*F*	*P-value*	*F critical*
Buffer time (BT)	18	4	573.183	1.37 E-60	2.479
Planning time (PT)	18	4	645.202	1.06 E-62	2.479

4.5 Conclusions

This research evaluated the LOS of the road hierarchy with snow effects integration from TTR point of view using ETC2.0 probe data. When Compared to traditional evaluations using ordinary travel time, the purposed method provided a detailed evaluation process by obtaining the comparative analysis between winter and non-winter seasons and also road hierarchy configuration. It considered and evaluated the changes of travel time distribution, the variation of travel time in different times of day, and the impact of snowfall on TTR by generated reliability indexes where BT and PT are the time reliability indicators. Regarding the change in TTR on the functional road system and weather fluctuations (snowfall), illustrations have simplified degrees of changes among road hierarchy system. However, the TTR of high standard highways such as expressway and national highway tended to change less and hardly had any effects. The TTR of the ordinary city roads such as prefectural, municipal, and community roads seemed to deteriorate while snow event occurred because of snow removal operation is the priority, and traffic demanded differentiations on the road hierarchy.

As for the major finding in the future, firstly, the explanation of travel time on different road hierarchy, the distribution of travel time used to clarify the phenomenal, expressway presented speedy characteristic of road hierarchy divergent others road hierarchy should be clarified. Although the high standard highway obtained the small amount of travel time, the portion of speedy users during non-winter was greater than the winter period because of the influence of snowfall. The results of the experiment are consistent with reality, which reflected the TTR degrees on individual road hierarchy. Second, the change of hourly effect, the fluctuation of time reliability during rush hour periods, demonstration of morning and afternoon rush hours, notices travel plan interruption with the large scale of travel time should be studied. The expressway shows a slight effect in either winter or non-winter. The winter season has seen the larger fluctuation of travel time than the non-winter period.

The last, the influence of snow depth on road hierarchy increased in corresponded to the amount of daily snowfall; thus, the time reliability also varied in regard to snow events. According to the level of snow depth, most of the road hierarchy seemed to have a steady low level of snow scale which occurred at a high scale especially on the national highway, which suffered the most comparing other road hierarchy because its function to distribute traffic volume from the intercity highway to the ordinary city roads. The study carried out the possibility to arrive on time over traffic demand and weather conditions throughout the travel time reliability viewpoint. The significant weather forecasts might be a main character to driving travel demand while weather condition is changing.

4.6 References

[1] B. Mehran and H. Nakamura, “Application of travel time reliability for performance oriented operational planning of expressways,” J. Japan Soc. Civel Eng. Ser. D3(infrastructure Plan. Manag., vol. 67, no. 3, pp. 283–299, 2011.

[2] M. J. Koetse and P. Rietveld, “The impact of climate change and weather on transport: An overview of empirical findings,” Transp. Res. Part D Transp. Environ., vol. 14, no. 3, pp. 205–221, 2009.

[3] I. Tsapakis, T. Cheng, and A. Bolbol, “Impact of weather conditions on macroscopic urban travel times,” J. Transp. Geogr., vol. 28, pp. 204–211, Apr. 2013.

[4] Cosmetic Ingredient Review Expert Panel, “Final report of the safety assessment of methacrylate ester monomers used in nail enhancement products.,” Int. J. Toxicol., vol. 24 Suppl 5, no. 5_suppl, pp. 53–100, Sep. 2005.

[5] C. Carrion and D. Levinson, “Value of travel time reliability: A review of current evidence,” Transp. Res. Part A Policy Pract., vol. 46, no. 4, pp. 720–741, 2012.

[6] D. Deka, “Factors associated with disability paratransit’s travel time reliability,” J. Transp. Geogr., vol. 48, pp. 96–104, 2015.

[7] U.S. Department of Transportation and Federal Highway Administration, “Travel Time Reliability: Making It There on Time, All the Time,” Fed. Highw. Adm. USA, pp. 1–18, 2006.

[8] K. Lyman and R. Bertini, “Using Travel Time Reliability Measures to Improve Regional Transportation Planning and Operations,” Transp. Res. Rec. J. Transp. Res. Board, vol. 2046, pp. 1–10, 2008.

[9] J. Bates, J. Polak, P. Jones, and A. Cook, “The valuation of reliability for personal travel,” Transp. Res. Part E Logist. Transp. Rev., vol. 37, no. 2–3, pp. 191–229, Apr. 2001.

[10] B. Mehran and H. Nakamura, “Implementing Travel Time Reliability for Evaluation of Congestion Relief Schemes on Expressways,” Transp. Res. Rec. J. Transp. Res. Board, vol. 2124, no. 1, pp. 137–147, 2009.

[11] A. Mondschein and B. D. Taylor, “Is traffic congestion overrated? Examining the highly variable effects of congestion on travel and accessibility,” J. Transp. Geogr., vol. 64, no. October 2016, pp. 65–76, 2017.

[12] H. K. Wong and J. M. Sussman, “Dynamic travel time estimation on highway networks,” Transp. Res., vol. 7, no. 4, pp. 355–370, 1973.

[13] X. Cai, J. J. Lu, Y. Xing, C. Jiang, and W. Lu, “Analyzing Driving Risks of Roadway Traffic under Adverse Weather Conditions: In Case of Rain Day,” Procedia - Soc. Behav. Sci., vol. 96, no. Cictp, pp. 2563–2571, 2013.

[14] V. Torrisi, M. Ignaccolo, and G. Inturri, “Estimating travel time reliability in urban areas through a dynamic simulation model,” Transp. Res. Procedia, vol. 27, pp. 857–864, 2017.

[15] D. Akin, V. P. Sisiopiku, and A. Skabardonis, “Impacts of weather on traffic flow characteristics of urban freeways in Istanbul,” Procedia - Soc. Behav. Sci., vol. 16, pp. 89–99, 2011.

[16] A. T. Ibrahim and F. L. Hall, “Effect of adverse weather conditions on speed-flow-

occupancy relationships," Transp. Res. Rec. J. Transp. Res. Board, vol. 1457, pp. 184–191, 1994.

[17] W. Yuan-Qing and L. Jing, "Study of Rainfall Impacts on Freeway Traffic Flow Characteristics," Transp. Res. Procedia, vol. 25, pp. 1533–1543, 2017.

[18] S. Datla, P. Sahu, H.-J. Roh, and S. Sharma, "A Comprehensive Analysis of the Association of Highway Traffic with Winter Weather Conditions," Procedia - Soc. Behav. Sci., vol. 104, pp. 497–506, 2013.

[19] N.-E. E. Faouzi, R. Billot, and S. Bouzebda, "Motorway travel time prediction based on toll data and weather effect integration," IET Intell. Transp. Syst., vol. 4, no. 4, p. 338, 2010.

[20] W. Brillon, J. Geistefeldt, and M. Regler, "Reliability of freeway traffic flow: a stochastic concept of capacity," in Proceedings of the 16th International Symphosium on Transportation and Traffic Theory, no. July, 2005, pp. 125–144.

[21] "National road in Hokkaido." [Online]. Available: https://www.hkd.mlit.go.jp/ky/kn/dou_kei/ud49g70000002ta4.html. [Accessed: 10-May-2018].

[22] "Japan Meteorological Agency." [Online]. Available: http://www.jma.go.jp/jma/indexe.html. [Accessed: 20-Apr-2018].

[23] D. Woodard, G. Nogin, P. Koch, D. Racz, M. Goldszmidt, and E. Horvitz, "Predicting travel time reliability using mobile phone GPS data," Transp. Res. Part C Emerg. Technol., vol. 75, pp. 30–44, 2017.

[24] M. Kyte, Z. Khatib, P. Shannon, and F. Kitchener, "Effect of Weather on Free-Flow Speed," Transp. Res. Rec. J. Transp. Res. Board, vol. 1776, no. 01, pp. 60–68, 2001.

[25] B. L. Smith, K. G. Byrne, R. B. Copperman, S. M. Hennessy, and N. J. Goodall, "An Investigation into the impact of rainfall," 83rd Annu. Meet. Transp. Res. Board, p. 15, 2004.

[26] B. Mehran, P. Student, and H. Nakamura, "PERFORMANCE EVALUATION OF HIGHWAY SEGMENTS USING TRAVEL TIME BASED PERFORMANCE MEASURES," 2007.

Chapter 5

Conclusions and Recommendations

5.1 Conclusions

A comprehensive study aims to augment the efficiency of intelligent data implements to urban sustainable development approach. Since the problems and research gaps have been found, this dissertation was divided into three sections which represented in chapter 2 – chapter 4.

5.1.1 Urban behavior observation using association analysis

Since the research have been proposed the chapter 2 – chapter 3 were established alighting with similar methodology, association rule mining. An association rule mining-based exploration of travel patterns in wide tourism areas using a Wi-Fi package sensing survey. It was first stage of research described in chapter 2, which was starting point of experiment of market basket analysis. The market basket analysis is a comprehensive reference in business and marketing research, its concept fundamental to identifying the next purchase from a ton of transactions database. We aim to adapt this method to solve the destination and tourism problems which recognized in chapter 3. When the first stage provided the appropriate results, working on the wireless probe data with association rule mining method, it was possible and accurate.

The implementation of association rules mining to a brunch of transaction dataset, which collection by using a Wi-Fi package sensor. Its captured intercommunication data of traveler on specific locations. Probe requests, it contains specific information of the individual user such as mac-address, time stamps, coordination, site, etc. Then, the algorithm called Apriori applied for frequent itemset mining and association rule mining over relational databases. In this position we acknowledged the wireless probe data has a potential in mobility study, because it compromised with similar conceptual work of "IF and THEN" statement. Besides, the adjustment of indices for a proper degree in the Apriori algorithm is a necessary procedure to get significant results.

5.1.2 Tourism and destination development point of view

The analysis involved the combination of several techniques, such as time series of visiting patterns, association mining implementation, and destination mining enhancement. These approaches were used to explore sequential patterns related to visiting attractive destinations. Moreover, the media access control number from enabled wireless communication devices was a key factor used to trace traveler movement in the targeted area. The application of association rule mining revealed the rules of destination visits. These rules depict strong relationships in terms of visit frequency, and they illustrate "if and then" statements for visitor movement analysis. These analytical methods are applicable to other tourist destination studies. The results show that, for weekend travel, factors such as seasonality tourism, distance to the location, and activities at the destination affect the type of destination selected and the number of destinations visited per trip.

Attractive destinations mining towards massive tourism area sustainable development on Wi-Fi tracking data in Chapter 3. The research pays more attention on destination management in sustainable tourism positions. However, there were rise some critical concerns that mention in the recommendations section. We found that the intelligent data such a probe requests data is suitable for use in mobility study, its ability to be the proper data source for location determination analyses, which pointed out the attractive locations in large-area study. Likewise, the association rules mining method appropriated in determining the pattern of visiting spots.

5.1.3 ETC2.0 probe data obtain travel time observation

Assessment of the impact of snowfall on travel time reliability considering different road types using ETC2.0 probe data in Chapter 4. Since the intelligent transport system rules the modern world of transportation management, tons of dataset generated thought leaps of technology. It used for managing and developing a better support system in the transport area and urban development. The study shows the appreciation of trajectory data which generating from the electronic toll collecting system in the highway network; it can use for travel time observation over the road hierarchy. While the significant finding, the evidence of travel time on different road hierarchy, the distribution of travel time used to clarify the phenomenal, expressway presented speedy characteristics of road hierarchy different others road hierarchy should be defined. Although the high standard highway obtained the small amount of travel time, the portion of speedy users during non-winter was more significant than the winter period because of the influence of snowfall. The results of the experiment are consistent with reality, which reflected the travel reliability degrees on individual road hierarchy. Hence, there is confirmation of the intelligent transport system and wireless communication to develop a new frontier of transport study. It practical and more efficient rather than traditional data collecting processes.

5.1.4 Advantage of intelligent wireless technology

There were several instruments that represented the intelligent technology for urban mobility observations, Wi-Fi package sensing devise and electronic toll systems accordingly. The results proof capability of them which generated and provided efficiency data collection for mobility exploration. According to the Wi-Fi probe data which generated for a small sensor device equipped in major tourist area, it used as a supreme ingredient for set up association analysis, which is a market research area application. It bought a modern tactical method for tourism and destination management study. Eventually, the ETC2.0 probe data will enable the implementation of travel time reliability evaluation as proposed methodology and integration of weather effect and travel time to make possible to develop an advanced traffic management system with respect to weather variability on road hierarchy, and to distribute an alternative support mechanism for the road operator’s decision making. The results bought the possibility to arrive on time over traffic demand and weather conditions on the pilot section throughout TTR viewpoint where the testbed site is Sapporo metropolitan. As mentions, the advantage of using technology in urban mobility observation is possible and practical. We can take an advantage from their ability with reliability, budget control and amount of data providing over time when compare to traditional survey method.

5.1.5 Sustainability assessment and perspectives

The emphasis of this study has been proposing sustainability in urban development. According to environmental concerns and community mobility improvement, the sustainable philosophy intended to be the core subject of the research. Understanding of traveler behavior by using technology integration, it provided an opportunity for mobility studies and city planning also open challenging to the modern era to observe people's movement and their activities in the vast area. Therefore, to proposing the sustainability context, it must be considered encouraging people to shift to public transportation rather than private transport. Moreover, the idea of how to maintain and improve urban mobility underneath the limitation of resources — the beginning of making a sustainable living atmosphere — we obligation understanding the behavior of travelers as much as possible also external influencing factors. Then essential details used as resources for policy and strategy planning process.

The perspective of sustainability. User's point of view, time management is the most important while considering mobility. Planning their trip and decision whether the best solution to be on time at the final destination, it's a principle of traveling conception. Regarding the results, travelers could take benefit from the attractive destinations identifying process, which is planning advantage on their trips such as route choice, season choice, destination choice, and depart and arrival time estimation. Thus, it contributed to sustaining urban mobility as progressive accordingly. Energy consumption and emission generated might decrease since travelers can plan their trips significantly, which is an eco-friendly contribution. Public transport might consider as primary mode choice when travelers would like to travel during rush hour or high season, the distance to the destination is one of a crucial factor that makes mode choice decision. Because travel time and travel costs are an essential factor in the travel plan, public transport likely providing more efficiency compare to private mode. These kinds of activities can make a change based on proper decision-making regarding travel time and energy consumption reduction. Traffic congestions decreasing, aligned with the low number of traffic accidents. Its strength impact the community in a macro-scale perspective.

5.1.6 Policy Implementation and guidelines

Planner's point of view. This study represented an essential methodology; It would be considered as guidance for urban planning and development to achieve the sustainability goal. Public policy and strategy are crucial mechanisms to compose the direction of city development. Thus, it is necessary to integrate the multi-layers of knowledge towards sustainability concepts according to the results that point out the essential location in the tourist area. Tourism management must be considered to policy and strategy-making process — the idea of reducing the impact on tourism atmosphere and local residence, environment, and pollution concern. Public transport services must be improved, and the existing facility is the priority — consideration of increasing public transport users throughout the future. An attractive destination remarked, it represents the demand side toward traveling and travel behavior, which is the planner's responsibility to answer the supply side. Advanced Traveler Information System (ATIS) recommended as an essential technical support solution for traveler facilities. Hence, the planning authority would consider dimensions of the possibility of approaching sustainable urban development, not only policy and strategy, but technology integration is necessary.

Lastly, this study emphasis frameworks, adaptive applications, and intelligent probe data proposed concerning the primary objective. The main advantage of the research contributes to the proactive point of view for further sustainable social development, not a limitation to general users, policymakers, and authorities. But vary sectors such as tourism, business, transportation, urban plan, etc. could consider this study as guidance for their situations. Also, the results have been responding to the research questions; accordingly, sustainability is an internationally intending topic.

5.2 Recommendations

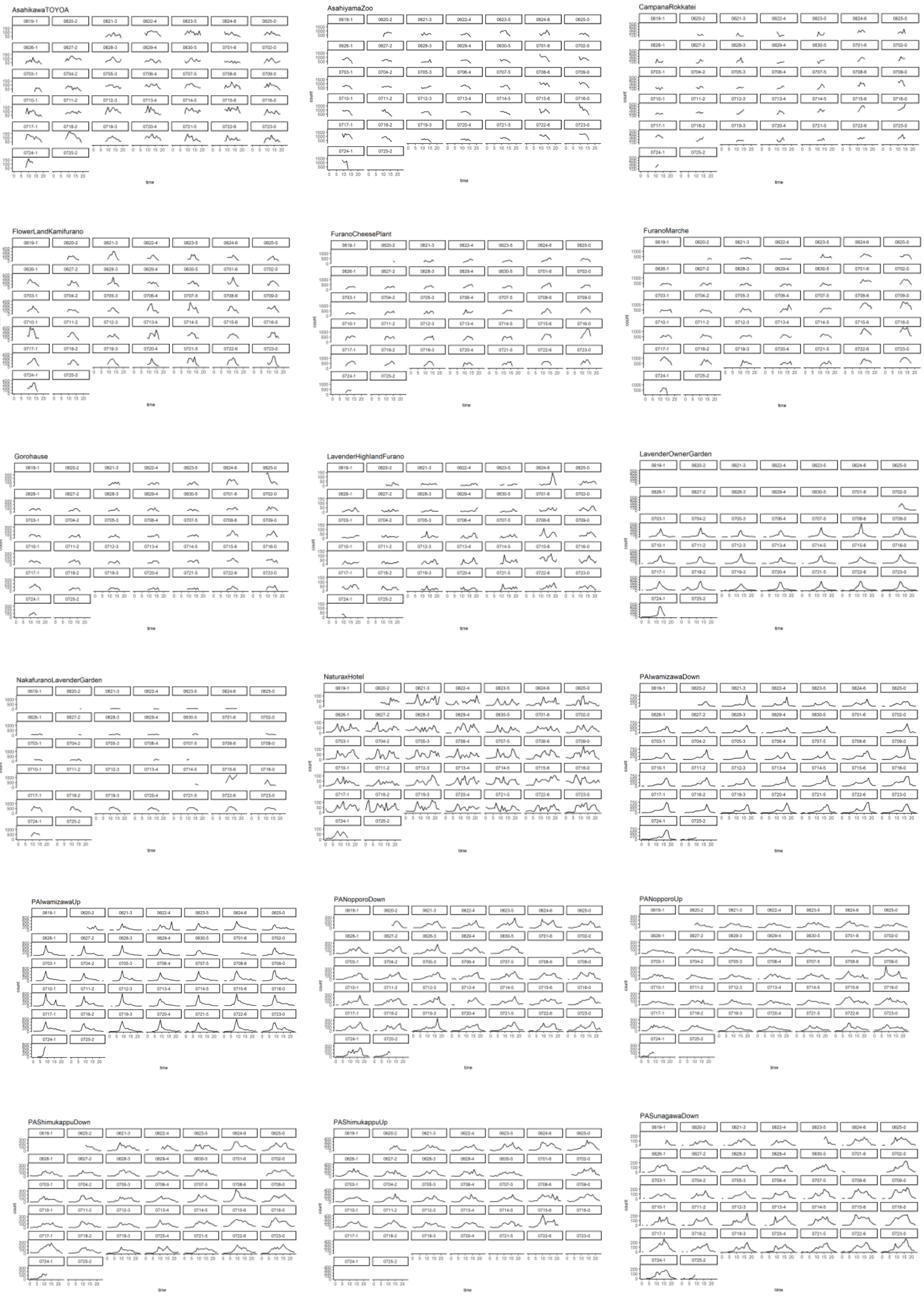
This dissertation expects that the recommendation to be presented in this section advantageous for future research, especially in the fields relating to tourism destination management and weather impact on travel time using intelligent data. The research separated into three sections; hence, the recommendation suggested following each chapter. The recommendations of this dissertation are as follows.

Chapter 2 and Chapter3: Since the survey adopted Wi-Fi scanner technology to collect on-site traveler movements, the advantage of using technology was declared. The long-term operation can be operated with a few human resources, also significantly in operation cost saving and gained a certain quality of data, when compared to conventional survey methods. However, there were some considered issues on the device limitation as working radius, and multiple devices installation may need in large scale and obstructed areas. In term of energy consumption, the operation device was consuming powerless from several sources such as portable batteries and USB charger, but long-term operation needs a secure energy source. The alternative source may consider, such as solar energy integration. Moreover, it is an off-line configuration device; real-time monitoring is incapable. Trace data stored in a tiny memory card and storage capability depends on its capacity. The future study, authors, consider expanding the range of Wi-Fi packet sensing devices utilization, to collecting mobility data in various scale area. It might consider expanding the range of the employed scanner to perform experiments inside specific tourism areas, such as farms, villages, or shopping malls, which cover more attractive destinations in the same area. We recommend expanding the field data collection period longer to get the more appropriate data from personal mobile device intercommunication on site as well as influencing factors such as seasonality and calendar events.

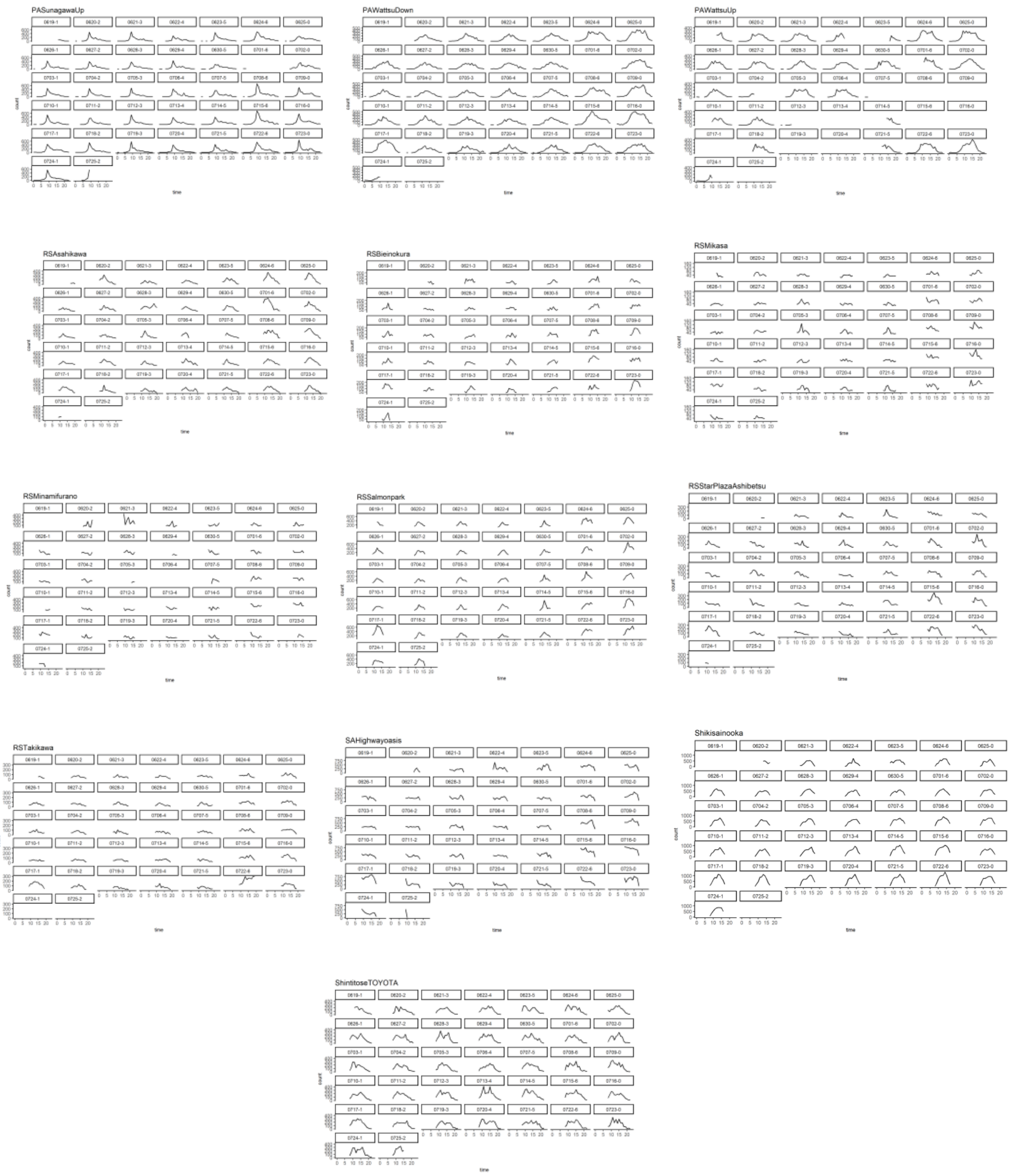
Chapter 4: The extension of this work will look forward to the analytics effort of conventional data and ETC2.0 probe data while considering the velocity and traffic demand affected by various weather conditions. The experiment on micro and macro scales of road network are also necessary as well as different detectors technology efficiency comparison.

Appendix A

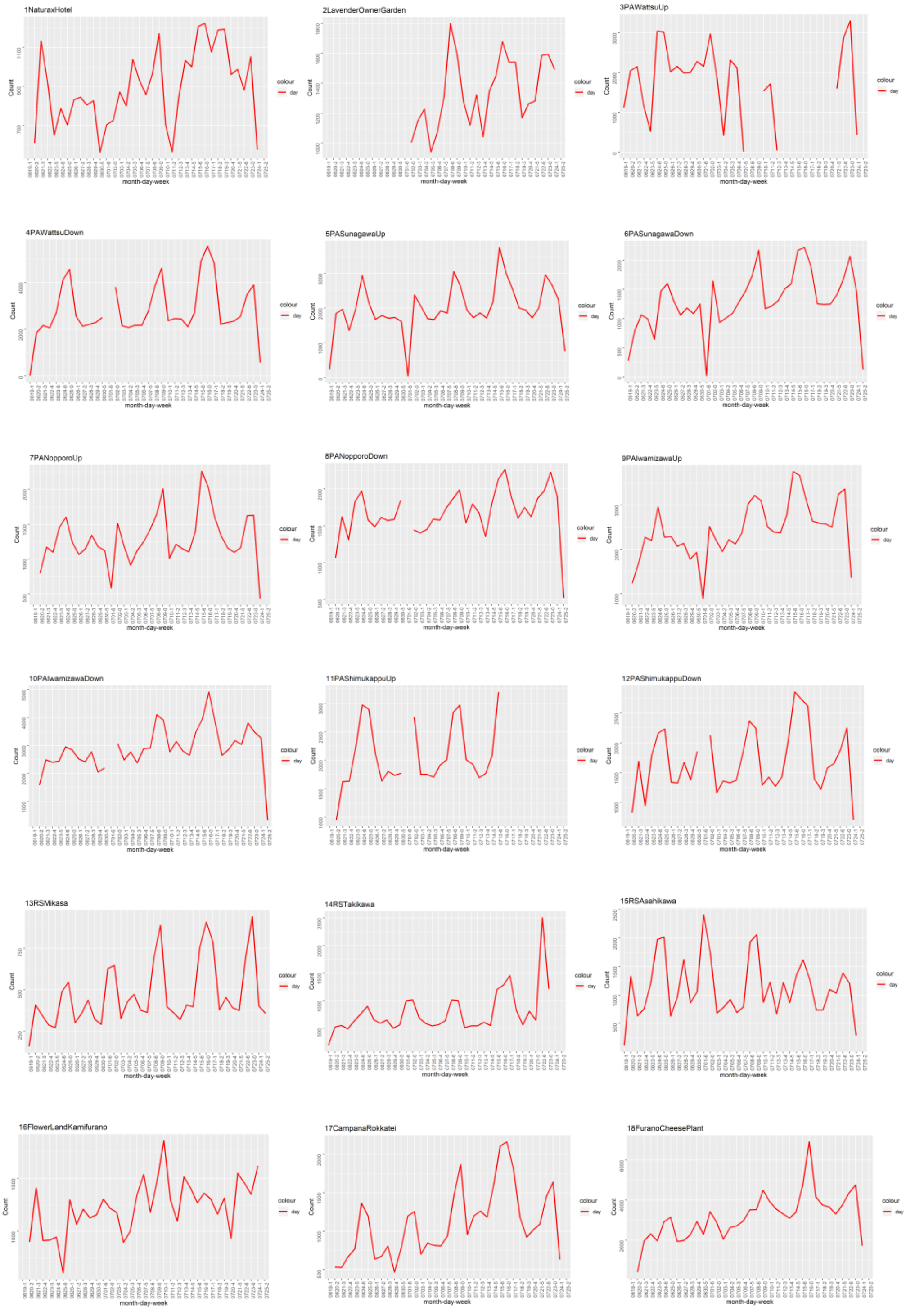
Additional Figures and Data


AsahikawaTOYOA
AsahiyamaZoo
CampanaRokkatei
FlowerLandKamifurano
FuranoCheesePlant
FuranoMarche
Gorohause
LavenderHighlandFurano
LavenderOwnerGarden
NakafuranoLavenderGarden
NaturaxHotel
PAIwamizawaDown
PAIwamizawaUp
PANopporoDown
PANopporoUp
PAShimukappuDown
PAShimukappuUp
PASunagawaDown

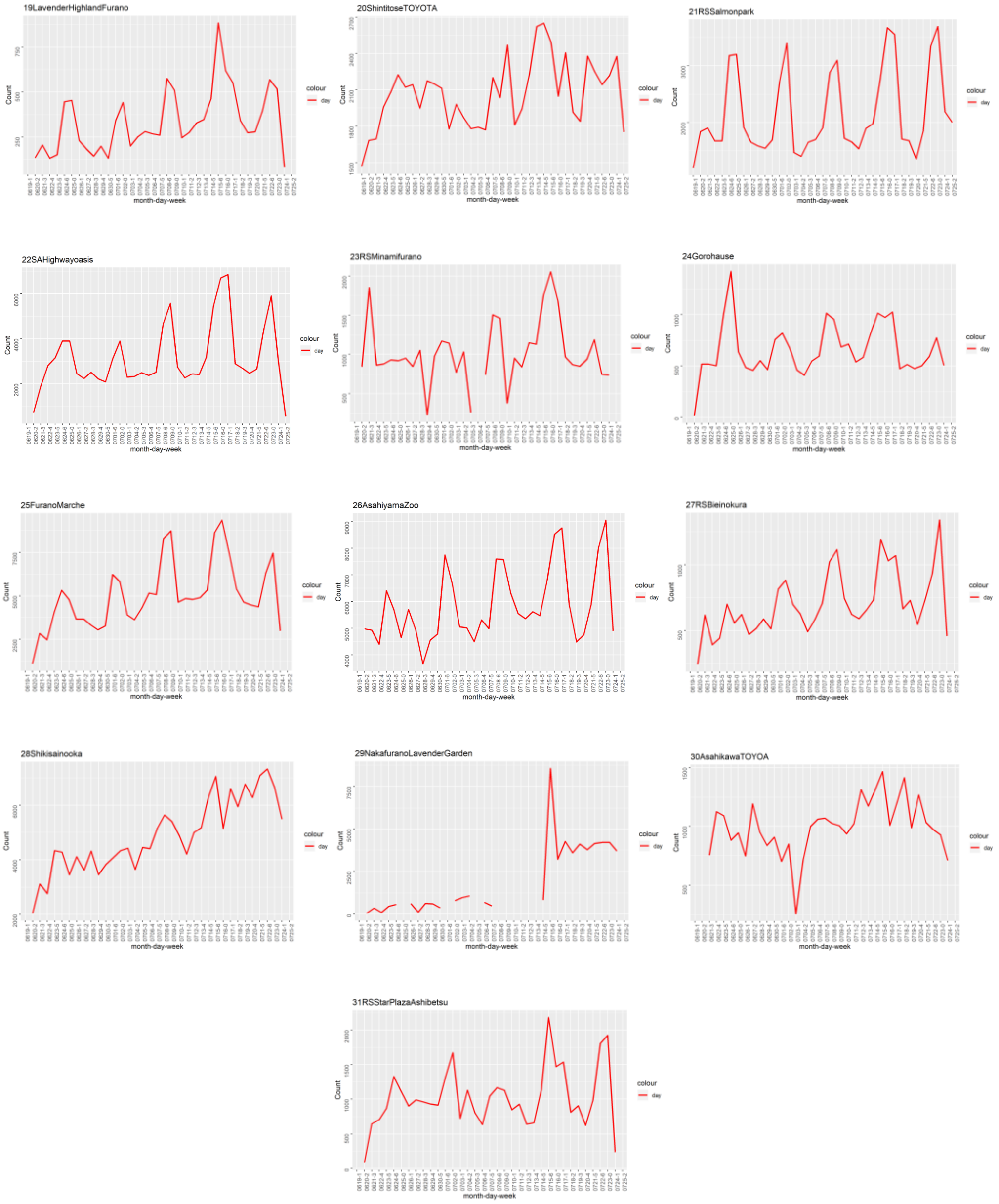
Appendix A1



Appendix A1 (continue)


1NaturaxHotel
2LavenderOwnerGarden
3PAWattsuUp
4PAWattsuDown
5PASunagawaUp
6PASunagawaDown
7PANopporoUp
8PANopporoDown
9PAIwamizawaUp
10PAIwamizawaDown
11PAShimukappuUp
12PAShimukappuDown
13RSMikasa
14RSTakikawa
15RSAsahikawa
16FlowerLandKamifurano
17CampanaRokkatei
18FuranoCheesePlant
Count
month-day-week
colour
day

Appendix A2



Appendix A2 (continue)

Appendix B

Journal,
Proceeding and Conferences,
and Society Membership

Journals

[1] Arreeras, T., Arimura, M., Asada, T., & Arreeras, S. (2019). Association Rule Mining Tourist-Attractive Destinations for the Sustainable Development of a Large Tourism Area in Hokkaido Using Wi-Fi Tracking Data. Sustainability, 11(14), 3967.

[2] ARREERAS, T., ENDO, M., TAKAHASHI, H., ASADA, T., & ARIMURA, M. (2019). An Association Rule Mining-Based Exploration of Travel Patterns in Wide Tourism Areas using A Wi-Fi Package Sensing Survey. Journal of the Eastern Asia Society for Transportation Studies, 13, 1099-1113.

[3] OKUMURA, K., OGOSHI, T., ASADA, T., ARIMURA, M., & ARREERAS, T. (2019). RBF Network Assessing Development of Downscaling Method for The Hourly Population Inflow: Case Study of Sapporo City Area. Journal of the Eastern Asia Society for Transportation Studies, 13, 1180-1193.

[4] Arreeras, T., Chongutsah, S., Asada, T., & Arimura, A. (2019). Factors Affecting Mode Choice in Accessing Railway Station Study in Nakhon Ratchasima. Transportation Research Procedia, (In press)

Proceedings and Conferences

[1] Arreeras, T., Arimura, M., Asada, T., & Arreeras, S. (2019). Attractive Destinations Mining Towards Massive Tourism Area Sustainable Development on Wi-Fi Tracking Data. In proceeding of The 10th International Seminar on Urban Transport, Tourism and Travel Behavior analysis, Sapporo, Japan, August 22-25, 2019, pp 151-158.

[2] Arreeras, T., Asada, T., & Arimura, M. (2018). A Study on Evaluation of Travel Time Reliability by Road Hierarchy in Winter using ETC2.0 Probe Data. The 10th Highway Data Meeting, Hiroshima, Japan, September 29-30, 2018.

[3] Arreeras, T., Kanbe, M., Asada, T., & Arimura, M. (2018). Analysis of Travel Time Reliability on Road Hierarchy in Winter Traffic Conditions using ETC2.0 Probe Data. In proceeding of the 16th ITS Asia-Pacific Forum FUKUOKA 2018, Fukuoka, Japan, May 8-10, 2018, pp 0156-0167.

[4] Thanasupsin, K., & Arreeras, T. (2007). Perceptions of Train, Intercity Buses and Low Cost Airline Travelers in Thailand. The Thailand Research Fund Master Research Grants Congress (TRF-MAG Congress), Jomtien Palm Beach Resort Pattaya at Chonburi, Thailand. April 20-22, 2007. (Thai)

[5] Thanasupsin, K., & Arreeras, T. (2006). The study of behavior and satisfaction on low cost airlines and intercity buses. The 11th National convention on civil engineering at Phuket, Thailand. April 20-22, 2006. (Thai)

[6] Arreeras, T., & Thanasupsin, K. (2006). A study of satisfactions of low cost airline travelers on factors affecting mode choice selection. The 3rd National convention on Transportation Engineering at Khon Kean, Thailand. November 24, 2006. (Thai)

Society Memberships

[1] World Conference on Transport Research Society (WCTRS)

[2] Eastern Asia Society for Transportation Studies (EASTS)

[3] The Engineering Institute of Thailand (EIT)