

Article

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

Tosporn Arreeras , Mikiharu Arimura \*, Takumi Asada and Saharat Arreeras

Division of Sustainable and Environmental Engineering, Muroran Institute of Technology, Muroran 050-8585, Japan; 17096008@mmm.muroran-it.ac.jp (T.A.); asada@mmm.muroran-it.ac.jp (T.A.); 19096002@mmm.muroran-it.ac.jp (S.A.)

\* Correspondence: arimura@mmm.muroran-it.ac.jp

Received: 22 May 2019; Accepted: 19 July 2019; Published: 22 July 2019



**Abstract:** The rise of radiofrequency scanner technology has led to its potential application in the observation of people's movements. 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.

**Keywords:** destination; Wi-Fi tracking data; massive area; tourist movement; survey technology; movement patterns; association rule mining; sustainability development

## 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<sup>2</sup> and encompasses unique terrains, alpinism, 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 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.

## 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.

### 2.1. Digital Footprints 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 user name 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 disembarking 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].

### 2.2. Market Basket Analysis in Tourism Studies

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].

### 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].

#### 2.4. Destination Types 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 1. 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].

**Table 1.** Type of destinations, ideal customer, and activities undertaken [35].

Type of Destination	Target Market	Activities
CBD <sup>1</sup> , 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

<sup>1</sup> Central business district.

### 3. Methods

#### 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 m 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 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 1. Example of probe request data from Wi-Fi scanner devices.

### 3.2. Data Collection 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 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 2.

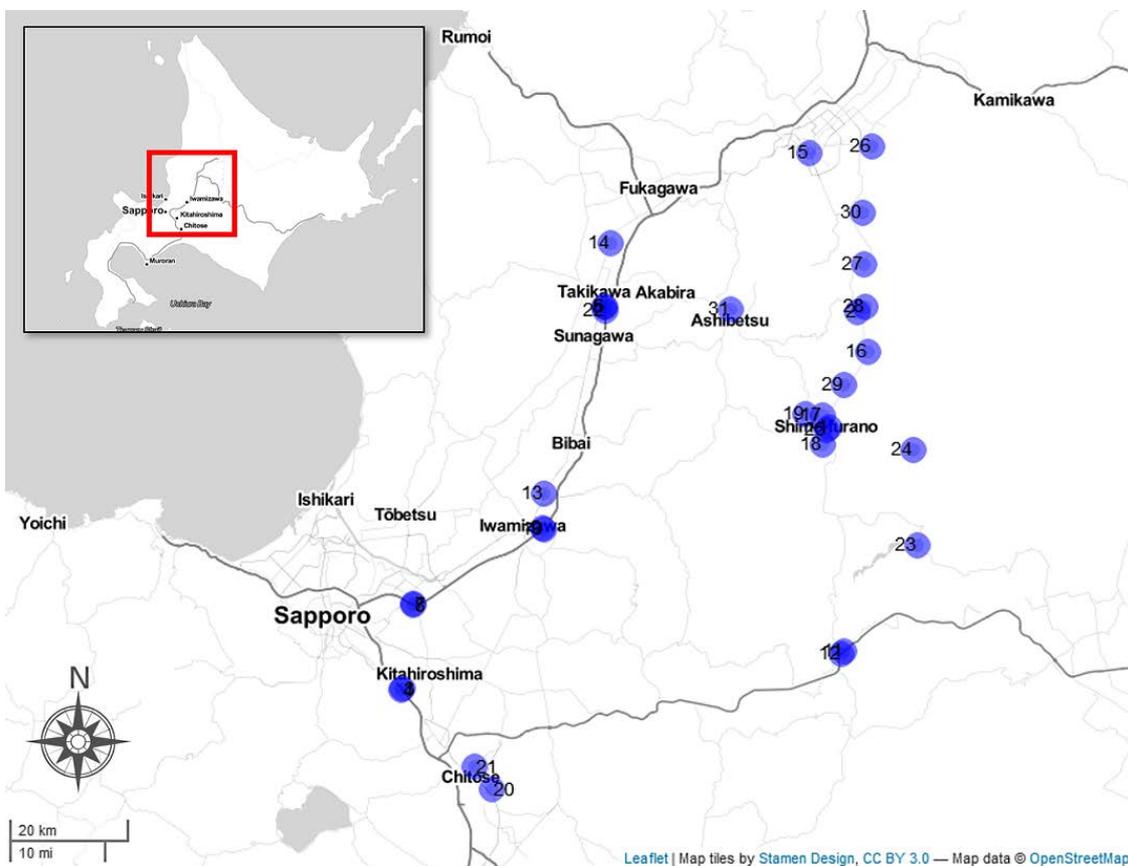


Figure 2. The coverage of observation locations of attractive destinations in Hokkaido.

**Table 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. 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.

The data cleansing process was carried out in several steps. (1) *Data integration of multi-location sensors*: Although the reception coverage range of the used Wi-Fi packet sensor is 300–400 m, the reception of probe requests from the user may be difficult because of the surrounding environment, such as buildings, crowds, weather, and so on. For that reason, multiple sensor installations were required. Then, the received data at the plurality of locations were combined to obtain the aggregate data for a particular site. (2) *Data extraction of arrival at and departure from sites*: The original data consists of continuously received records that were captured at intervals of several seconds. Except for the recording at the time of initial reception (arrival) and final reception (departure) at a site, the recordings were unnecessary and excluded. (3) *Exclusion of data from MAC addresses detected for ten days or more*: Arrival and departure data were arranged in chronological order for each MAC address. The round-trip time of the user was obtained by taking the difference between the arrival at the first site and departure from the last site. A stay of more than ten days is likely to be a resident or a user for business purposes, so such data were excluded.

### 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 (1). The confidence of a rule represents the accuracy or conditional probability of the itemset appearing, and it is given in Equation (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.

$$\text{Support}(X \Rightarrow Y) = \frac{\text{Number of transactions with both } X \text{ and } Y}{\text{Total number of transactions}} = P(X \cap Y) \quad (1)$$

$$\text{Confidence}(X \Rightarrow Y) = \frac{\text{Number of transactions with both } X \text{ and } Y}{\text{Total number of transactions } X} = \frac{P(X \cap Y)}{P(X)} \quad (2)$$

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) and (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].

$$\text{Expected Confidence}(X \Rightarrow Y) = \frac{\text{Number of transactions with } Y}{\text{Total number of transactions}} = P(Y) \quad (3)$$

$$\text{Lift}(X \Rightarrow Y) = \frac{\text{Confidence}(X \Rightarrow Y)}{\text{Expected Confidence}(X \Rightarrow Y)} = \frac{P(X \cap Y)}{P(X) \cdot P(Y)} \quad (4)$$

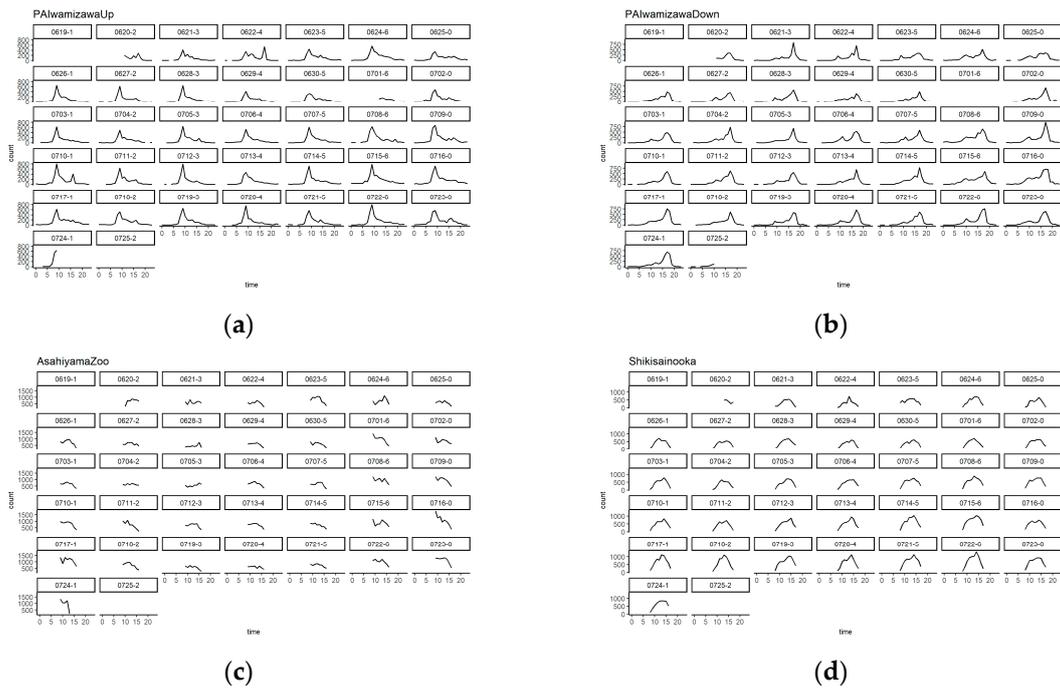
The ARM method has been used not only for market research but also for tourism, urban development, and various other fields of study. For example, ARM has been applied in areas such as sequential pattern mining analysis to extract travel patterns in transportation research [17]; it has also been used 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.

## 4. Results

### 4.1. Visiting Patterns in Time Series

#### 4.1.1. Hourly Patterns

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 3a,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.

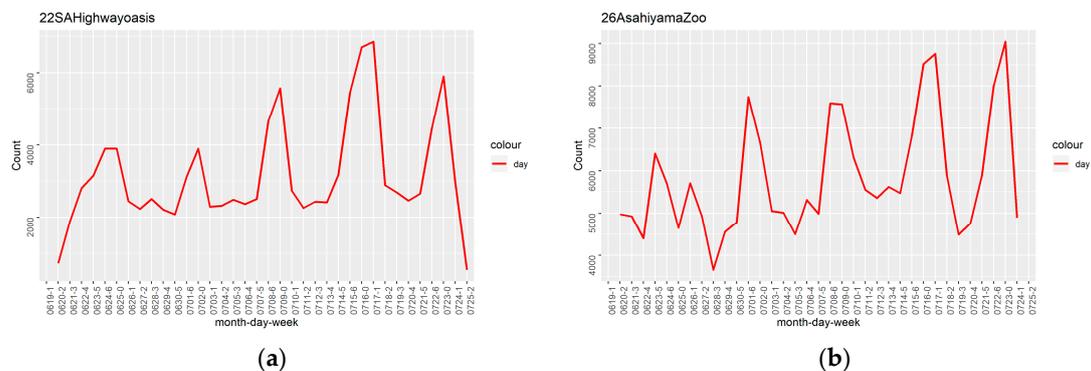


**Figure 3.** Hourly patterns of service and parking stations: (a) Inbound; (b) outbound. Hourly patterns at attractive tourist destinations: (c) The zoo; (d) flower garden.

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 3c,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.

#### 4.1.2. Weekly Patterns

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 4a) and an attractive destination (see Figure 4b) reveal similar results, with a both showing a significant increase in visitors on weekends.



**Figure 4.** Comparison of weekly patterns between (a) the service and parking area SAHighwayoasis station and (b) the tourism destination AsahiyamaZoo.

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.

## 4.2. Implementation of Association Mining

### 4.2.1. 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 5 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. 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.

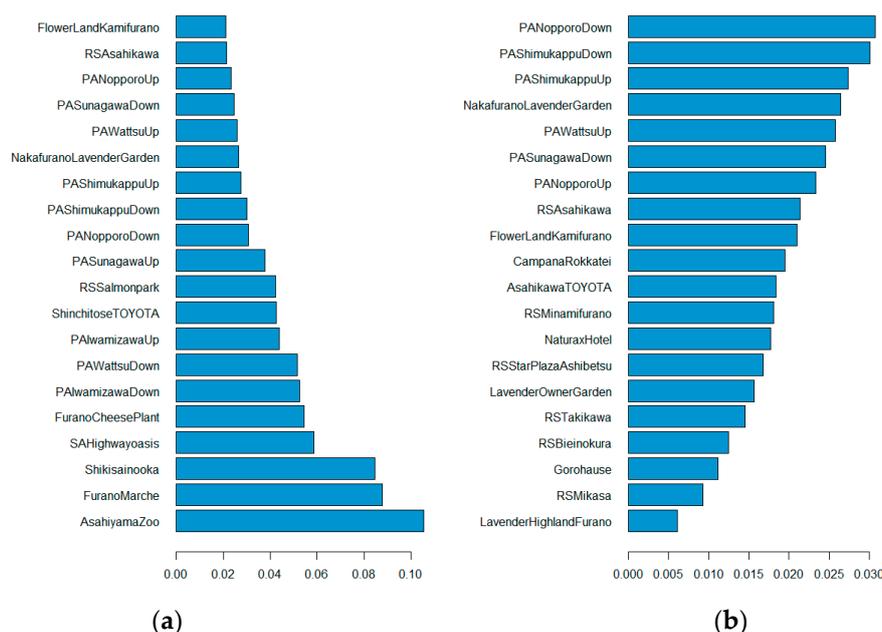


Figure 5. Relative distribution of visiting frequency—(a) 20 most visited sites; (b) 20 least visited sites.

### 4.2.2. Association Rule of Destination

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). The distribution of rule lengths indicates visitors’ preferences for how many locations they visit on their trip. Table 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 5.

**Table 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 4.** The distribution of rule lengths.

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

**Table 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

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.

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 6). 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.

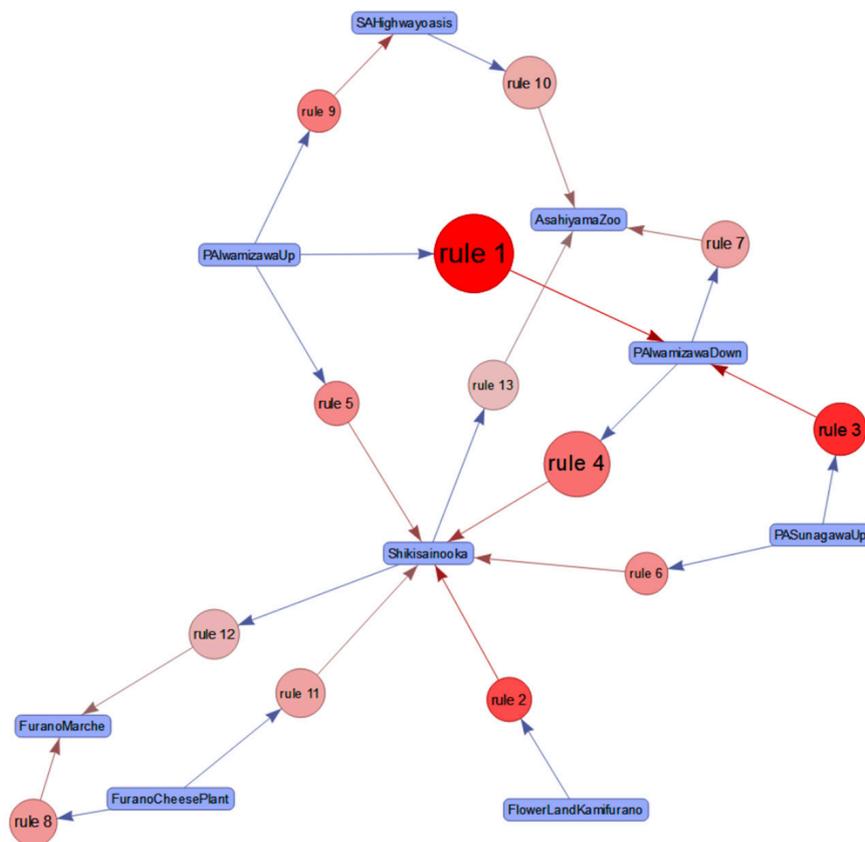


Figure 6. Visualization of all destination rules.

### 4.3. Destination Mining Enhancement

#### 4.3.1. Implementation of Statistical Hypothesis Test

The finalized association rules in Table 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 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 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 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

#### 4.3.2. 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 7 illustrates 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.

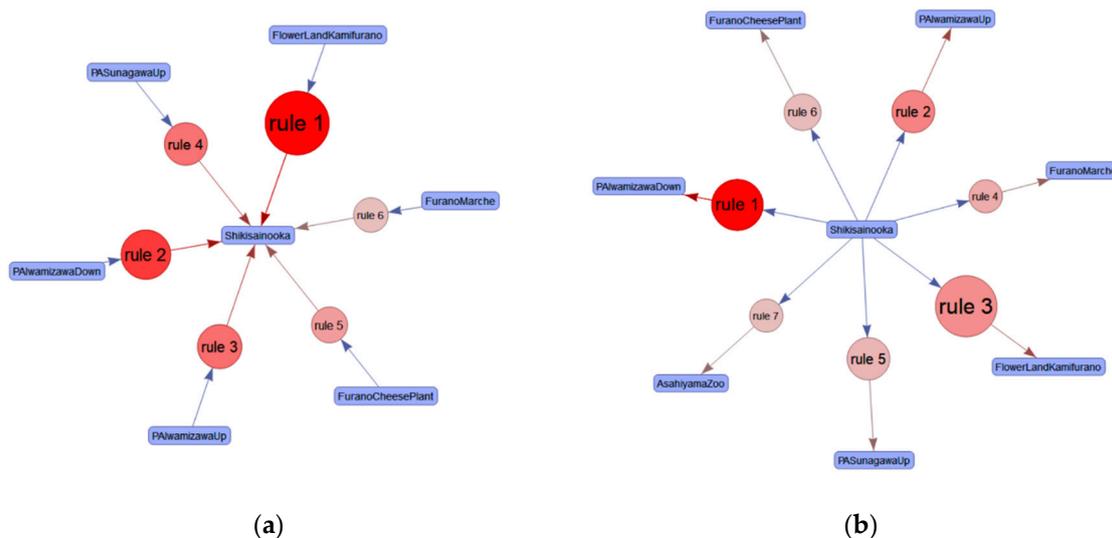


Figure 7. Recognized destination rules of Shikisainooka: (a) Inbound and (b) outbound visiting.

Table 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 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

### 5. Discussion and 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.

A future study might expand 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. Real-time monitoring might be necessary for detailed tourist mobility analysis, including the investigation of short-term and long-term visits as well as influencing factors such as seasonality and calendar events.

**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.

**Author Contributions:** T.A. (Tosporn Arreeras) and M.A. designed the research; field survey T.A. (Takumi Asada); analyzed data and writing the paper, T.A. (Tosporn Arreeras); cleaned data, T.A. (Tosporn Arreeras) and S.A.; supervision, T.A. (Takumi Asada) and M.A.

**Funding:** This research was funded by National Institute for Land and Infrastructure Management under the Ministry of Land, Infrastructure, Transport, and Tourism (Japan).

**Acknowledgments:** This research was conducted with the support from the National Institute for Land and Infrastructure Management under the Ministry of Land, Infrastructure, Transport, and Tourism (Japan) that contributes to improving the quality of road policy "R&D on data driven traffic demand and space management utilizing ETC 2.0 probe information."

**Conflicts of Interest:** The authors declare no conflict of interest.

## 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. [CrossRef]
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. [CrossRef]
3. Abreu Novais, M.; Ruhanen, L.; Arcodia, C. Destination competitiveness: A phenomenographic study. *Tour. Manag.* **2018**, *64*, 324–334. [CrossRef]
4. Mckercher, B.; Lau, G. Movement Patterns of Tourists within a Destination. *Tour. Geogr.* **2008**, *10*, 355–374. [CrossRef]
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 22 January 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. [CrossRef]
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. [CrossRef]
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. [CrossRef]
10. EMEL, G.G.; TAŞKIN, Ç.; AKAT, Ö. Profiling a Domestic Tourism Market by Means of Association Rule Mining. *Anatolia* **2007**, *18*, 334–342. [CrossRef]
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. [CrossRef]

12. Shoval, N.; Isaacson, M. Tracking tourists in the digital age. *Ann. Tour. Res.* **2007**, *34*, 141–159. [[CrossRef](#)]
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.; Piazzolo, 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. [[CrossRef](#)]
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. [[CrossRef](#)] [[PubMed](#)]
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. [[CrossRef](#)]
17. Bermingham, L.; Lee, I. Spatio-temporal sequential pattern mining for tourism sciences. *Procedia Comput. Sci.* **2014**, *29*, 379–389. [[CrossRef](#)]
18. Shoval, N.; Ahas, R. The use of tracking technologies in tourism research: The first decade. *Tour. Geogr.* **2016**, *18*, 587–606. [[CrossRef](#)]
19. Musa, A.B.M.; Eriksson, J. Tracking unmodified smartphones using wi-fi monitors. In Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems-SenSys'12, Toronto, ON, Canada, 6–9 November 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. [[CrossRef](#)]
21. Danalet, A.; Tinguely, L.; de Lapparent, M.; Bierlaire, M. Location choice with longitudinal WiFi data. *J. Choice Model.* **2016**, *18*, 1–17. [[CrossRef](#)]
22. Potorti, F.; Crivello, A.; Girolami, M.; Barsocchi, P.; Traficante, E. Localising crowds through Wi-Fi probes. *Ad Hoc Netw.* **2018**, *75*, 87–97. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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*, e88859. [[CrossRef](#)] [[PubMed](#)]
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, Jinan, China, 14–16 August 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. [[CrossRef](#)]
32. Mooney, C.H.; Roddick, J.F. Sequential pattern mining—approaches and algorithms. *ACM Comput. Surv.* **2013**, *45*, 1–39. [[CrossRef](#)]
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. [[CrossRef](#)]
34. Hwang, Y.-H.; Gretzel, U.; Fesenmaier, D.R. Multicity trip patterns. *Ann. Tour. Res.* **2006**, *33*, 1057–1078. [[CrossRef](#)]
35. Buhalis, D. Marketing the competitive destination of the future. *Tour. Manag.* **2000**, *21*, 97–116. [[CrossRef](#)]

36. Raun, J.; Ahas, R.; Tiru, M. Measuring tourism destinations using mobile tracking data. *Tour. Manag.* **2016**, *57*, 202–212. [[CrossRef](#)]
37. Crouch, G.I.; Ritchie, J.R.B. Tourism, Competitiveness, and Societal Prosperity. *J. Bus. Res.* **1999**, *44*, 137–152. [[CrossRef](#)]
38. Franzoni, S.; Bonera, M. How DMO Can Measure the Experiences of a Large Territory. *Sustainability* **2019**, *11*, 492. [[CrossRef](#)]
39. Ryan, C. Tourism and marketing—a symbiotic relationship? *Tour. Manag.* **1991**, *12*, 101–111. [[CrossRef](#)]
40. Gunn, C.A. *Tourism Planning: Basics, Concepts, Cases*, 3rd ed.; Taylor & Francis: Abingdon, UK, 1994; ISBN1 0844817430. ISBN2 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. [[CrossRef](#)]
42. Frechtling, D.C. *Forecasting Tourism Demand: Methods and Strategies*, 1st ed.; Butterworth-Heinemann: Great Britain, UK, 2001; ISBN 9781136402425.
43. Poon, A. *Tourism, Technology and Competitive Strategies*; CAB International: Wallingford, UK, 1993; ISBN 0851987516.
44. Gopikaraman, R.; Rameshkumar, T.; Senthil Kumaran, B.; Ilangoan, 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. [[CrossRef](#)]
46. North, M. *Data Mining for the Masses*; Global Text Project: Athens, GA, USA, 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. [[CrossRef](#)]
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 International Workshop on Applying Data Mining in e-Learning (ADML 2007), Crete, Greece, 18 September 2007; pp. 13–22.
49. Hahsler, M. arulesViz: Interactive Visualization of Association Rules with R. *R J.* **2017**, *9*, 163–175. [[CrossRef](#)]
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, Hubei, China, 25–26 September 2008; pp. 414–417.
52. Brin, S.; Motwani, R.; Silverstein, C. Beyond market baskets. *ACM Sigmod Rec.* **1997**, *26*, 265–276. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).