Characteristic Comparison of Tourist Areas Using the Statistical Information

Atsushi TANAKA [a], Naohisa OKAMOTO [b], Toshihiro SUZUKI [c],
Reiko ASANO [c], Yoji SHIRAKAWA [c]

[a] *Graduate School of Systems and Information Engineering of Tsukuba University, 1-1-1 Tennodai, Tsukuba-City, Ibaraki, 305-8577, Japan;*
[a] *E-mail: s1620476@sk.tsukuba.ac.jp*
[b] *Professor of Division of Policy and Planning Sciences of Tsukuba University, 1-1-1 Tennodai, Tsukuba-City, Ibaraki, 305-8577, Japan;*
[b] *E-mail: okamoto@sk.tsukuba.ac.jp*
[c] *DOCOMO Insight Marketing, .Inc , 2-5-5 Shinbashi, Minato-Ward, Tokyo, 105-0004, Japan;*
[c] *E-mail: youji.shirakawa.ge@dcm-im.com*

Abstract: In recent years, along with the improvement of the information and communication technology, mobile sensing data obtained by utilizing applications and information obtained from position data of mobile phones and other devices are provided by various companies in Japan. This research confirmed the demographic dynamics in tourist areas and compared the characteristics of tourism behaviors using "Mobile-Spatial-Statics" which is the statistical information provided by NTT DOCOMO, Inc. The target areas were classified by cluster analysis using "Daytime-Nighttime Population Ratio" and "Nightly Population Ratio". In addition, tourist areas were illustrated for each cluster, showing the tourism characteristics.

Keywords: Tourism, Statistical Information, Cluster Analysis, Japan

1. INTRODUCTION

In March 2016, the Japanese government formulated the "Tourism Vision to Support Japan of Tomorrow". It set new targets for the tourism developed country and raised the tourism reform. In Japan that has entered an era of declining population, especially in rural areas, thus expectations for tourism are great for the regional creation.

In recent years, along with the improvement of the information and communication technology (ICT), mobile sensing data obtained by utilizing applications and the location data of mobile phones and other devices are provided by various companies in Japan. Therefore, how to use these data in the tourism field is sought. "Mobile-Spatial-Statistics" provided by NTT DOCOMO, Inc. which is the statistical information on data generated when providing telecommunication services. Although this data is organized in prefectures, municipalities, and meshes, the amount of data is large and it is possible to grasp highly accurate demographics by time zone and day.

There are research focusing on the behavior of tourists using the GPS data. Furutani (2006) proposed the classification method of tourists' behaviors using individual attributes and their behavior dynamics obtained by GPS. De Cantis *et al.* (2016) analyzed behavior patterns of tourists traveling around the region from the port of Palermo, Italy. Also, Ubukata *et al.* (2013)

analyzed tourism dynamics for Ishikawa pref. by using the trajectory of about 800,000 cell phone users. Katagiri *et al.* (2015) grasped the activity area of foreign tourists based on the activity trajectories collected using GPS logger and examined the possibility of evacuation guidance. In addition, Japan Tourism Agency aimed at the tourism promotion utilizing tourism big data, conducted analysis using applications, base station information, SNS, etc. As described above, with the recent improvement of ICT, the possibility of tourists' behavior analysis has expanded.

In this research, we confirmed the population concentration and variability of the tourist areas and categorized tourist areas as day trip type and accommodation type. By using "Mobile-Spatial-Statistics", we grasped the population dynamics close to the actual situation, and confirmed the characteristics of tourist areas based on abundant data.

2. DATA OVERVIEW

"Mobile-Spatial-Statistics" estimated the population based on the information of base station accessed by DOCOMO's mobile phones, sample size is large, and data can be acquired even in rural areas. It is possible to use data that summarized mobile phones existing in the base station area for each prefecture, municipalities, and mesh. However, it is necessary to note that the continuous movements such as "from where to move to where" of individual mobile phones cannot be confirmed because of the data to the de-identification, aggregation, and anonymization processes. Table 1 shows the Overview of data used in this study. It is available weekends data of the first week of every month and the data is tabulated in municipalities unit and it is possible to use the age, sex, and residential area (municipalities unit) as attribute data.

Table 1. Overview of data in this research

Period	January - December, 2015
Target	Domestic residents who uses DOCOMO's mobile phone
Date	Weekends of the first week of the month
Target Area	Municipalities unit
Residence	Municipalities unit

3. DATA ANALYSIS

3.1 Analyzed areas

In this research, we selected the top 50 municipalities as areas to be analyzed, based on the attractiveness ranking of the "Regional Brand Investigation 2015" conducted by Brand Research Institute, Inc. As for this ranking, asking "How much do you feel attractiveness" of any region's name, evaluating the answer in 5 stages, and the score is calculated with "very attractive" as 100 points, "somewhat attractive" as 50 points, and the rest as 0 points. Figure 1 shows the location of the Analyzed areas and Table 2 shows analyzed areas and the peak month for each area. Incidentally, ordinance-designated cities can obtain data of Ward, but the representative Ward is selected, and the Ward name is omitted hereafter. The peak month is the

one with the maximum average of the population by the time zone of the first weekend on the first week. The peak month is May, followed by January and December, and in May or January

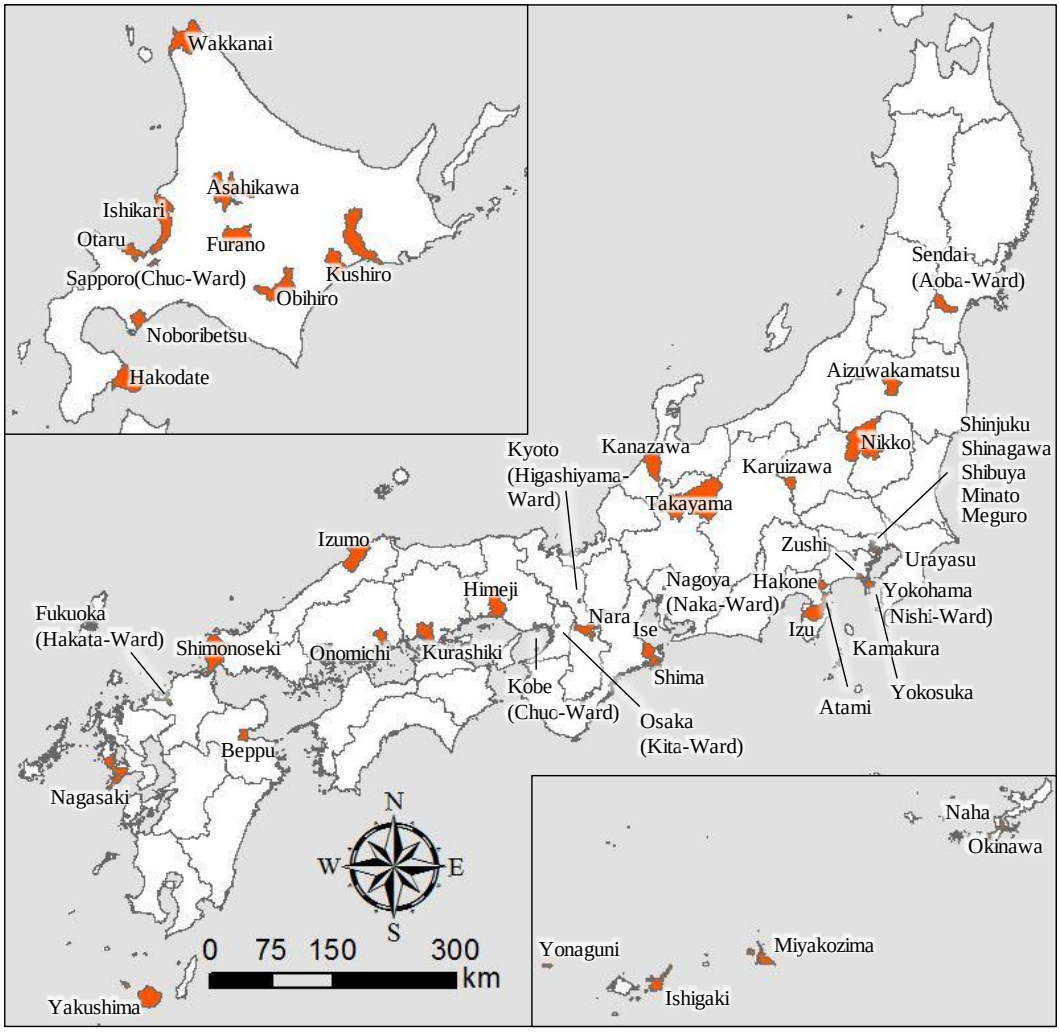

Figure 1. Analyzed areas

Table 2. Analyzed areas and their peak month

No.	Municipality Name	Peak Month	No.	Municipality Name	Peak Month	No.	Municipality Name	Peak Month
1	Hakodate	May.	21	Himeji	May.	41	Meguro	Apr.
2	Sapporo (Chuo-Ward)	Dec.	22	Shinjuku	Dec.	42	Kurashiki	Jan.
3	Kyoto (Higashiyama-Ward)	Apr.	23	Izu	May.	43	Obihiro	Jan.
4	Yokohama (Nishi-Ward)	Feb.	24	Izumo	Jan.	44	Ishikari	Aug
5	Otaru	May.	25	Ise	Jan.	45	Wakkanai	Jul.
6	Kobe (Chuo-Ward)	Dec.	26	Nagasaki	Mar.	46	Takayama	May.
7	Furano	May.	27	Kushiro	Jan.	47	Zushi	Aug
8	Kamakura	Jun.	28	Nara	Apr.	48	Yonaguni	Jul.
9	Kanazawa	Jun.	29	Shima	May.	49	Shimonoseki	Jan.
10	Karuizawa	May.	30	Hakone	May.	50	Aizuwakamatsu	May.
11	Beppu	May.	31	Nagoya (Naka-Ward)	Dec.			
12	Yakushima	May.	32	Urayasu	Dec.			
13	Naha	Dec.	33	Noboribetsu	Jan.			
14	Nikko	May.	34	Osaka (Kita-Ward)	Dec.			
15	Atami	May.	35	Shibuya	Nov.			
16	Ishigaki	Jan.	36	Shinagawa	Dec.			
17	Okinawa	Sep.	37	Miyakojima	Jan.			
18	Asahikawa	Jan.	38	Minato	Dec.			
19	Fukuoka (Hakata-Ward)	Apr.	39	Yokosuka	Apr.			
20	Sendai(Aoba-Ward)	Sep.	40	Onomichi	May.			

it tends to be populous because of a long holiday period.

3.2 Cluster Analysis

Subsequently, tourist areas were classified by cluster analysis using two statistics of “Daytime-Nighttime Population Ratio” and “Nightly Population Ratio”. “Daytime-Nighttime Population Ratio” is a value obtained by dividing the average value of the nighttime (from 22 o'clock to 5 o'clock) population by the average value of the daytime (from 6 o'clock to 21 o'clock) population. It is possible to grasp areas with large population dynamics during the day. “Nightly Population Ratio” is the average value of the nighttime population (0 o'clock to 5 o'clock) on Sunday divided by the average value of the nighttime population on Saturday (0 o'clock to 5 o'clock). By using this ratio, tourist areas can be classified as day trip type and accommodation type.

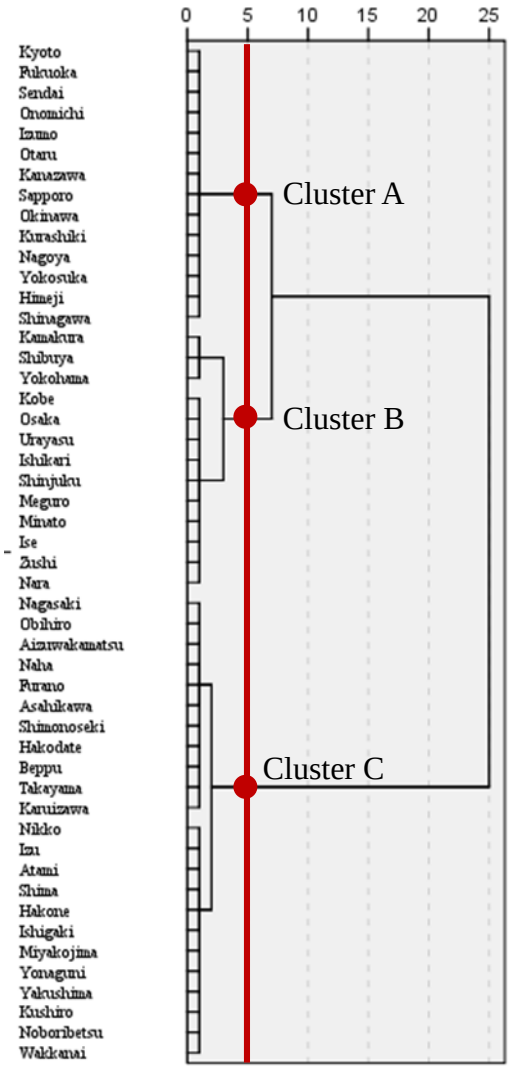

Figure 2. Result of cluster analysis

As a method of distance measurement in the analysis, the Ward method was used and the square Euclidean distance was adopted. The number of clusters was determined to 3 from the dendrogram (Figure 1). Significant results were obtained at a level within 5% by analysis of variance, with $F(2, 47) = 125.9$ ($p < 0.001$) for the “Daytime-Nighttime Population Ratio” and $F(2, 47) = 4.6$ ($p < 0.05$) for the “Nightly Population Ratio”. The three clusters are identified as A, B, and C. The characteristics of each cluster are as follows. Cluster A has areas with large population fluctuations in the day and night, Cluster B has cities in an intermediate position in both ratios. Cluster C includes cities with a large “Nightly Population Ratio”.

3.3 Population Distribution by Time Zone and Month

In this section, the target area is shown as an example, and its characteristics are compared to the population distribution by time zone and month. Here, we created population distribution charts, except for the data of people living in the target area. Figure 3 shows the legend of population distribution charts.

3.3.1 Cluster A

Clusters A includes cities such as Shibuya, Osaka, Urayasu and Ise and so on. As an example, we confirmed the population distribution of Shibuya, Urayasu and Ise. Shibuya-Ward, Tokyo (Figure 4) is known as a crowded shopping city for young people and is one of the secondary city centers. Looking at the population distribution chart, the population sharply increases after 6:00 am, then peaks around 15 o'clock, and the population decreases towards the night. As the daytime population is very large and the nighttime population is relatively small, thus it was

possible to read the dynamics affecting mainly the first and last train. Also, January is a long holiday period, many people visiting parents' house, and the population will be decreased. Furthermore, the data on Saturdays in November showed that the population at night is large due to the Halloween event.

Urayasu-City, Chiba (Figure 5) is an area with Tokyo Disney Resort (TDR), and many people visit from all over the country throughout the year. According to the figure, the population is large in the daytime time zone as a whole, and the movement affected by the opening hours is seen. Also, there was a difference in the number of visitors per month. This is thought to be related to the events in the park, the vacation, and the influence of the weather.

Ise-City, Mie (Figure 6) has a historical landscape preserved and many people visit Ise-Jingu-Shrine. It tends to reach a peak around 13 o'clock. Compared with the examples of Shibuya-

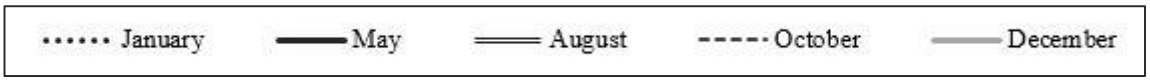

Figure 3. legend of population distribution charts

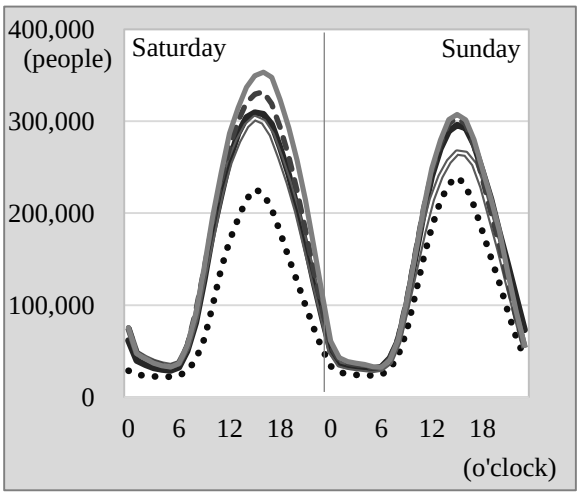

Figure 4. Population distribution chart **(**Shibuya-Ward, Tokyo**)**

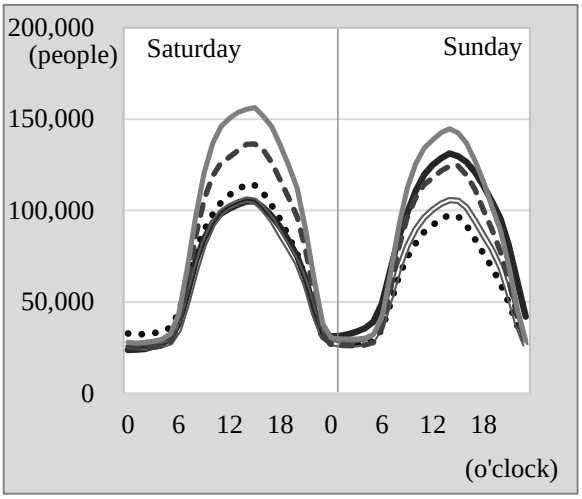

Figure 5. Population distribution chart **(**Urayasu-City, Chiba**)**

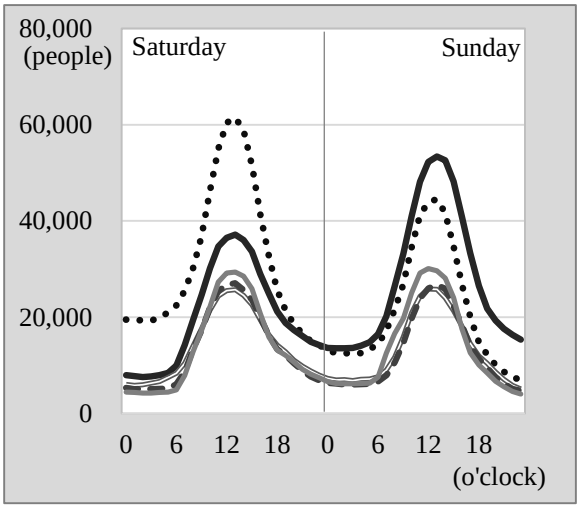

Figure 6. Population distribution chart **(**Ise-City, Mie**)**

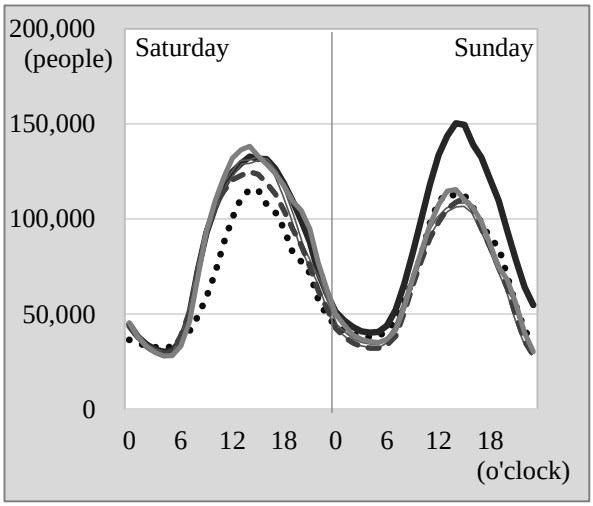

Figure 7. Population distribution chart **(**Fukuoka-City(Hakata-Ward), Fukuoka**)**

City and Urayasu-City as mentioned above, it seems that the population other than the peak time zone is small. In addition, In January, the population fluctuation is large because there are many visitors to Ise-Jingu-Shrine.

3.3.2 Cluster B

Cluster B includes local cities such as Sapporo, Fukuoka, Himeji and so on. As an example, confirmed the population distribution of Fukuoka and Himeji.

First, confirm the population distribution chart of Fukuoka-City (Hakata-Ward), Fukuoka (Figure 7). Fukuoka-City is the most prosperous commercial city in Kyushu, especially in Hakata-Ward there are many commercial facilities. The similar population dynamics can be seen in any months, but the population is large on Sundays in May and the population is somewhat less on Saturdays in January. Also, since the population movement is not constantly at night, it seems that the population moves slightly at night.

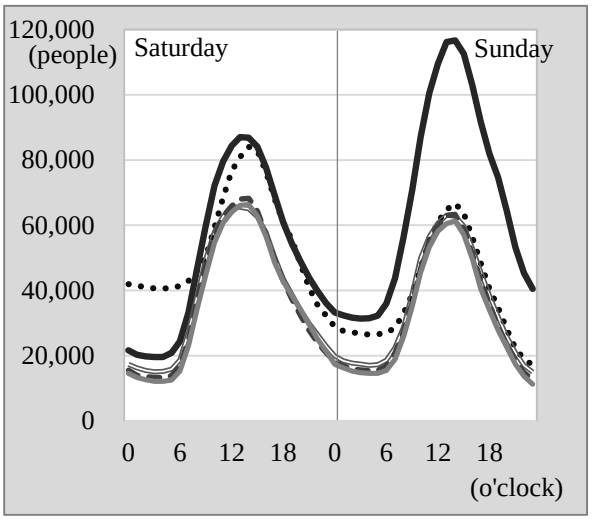

Figure 8. Population distribution chart **(**Himeji-City, Hyogo**)**

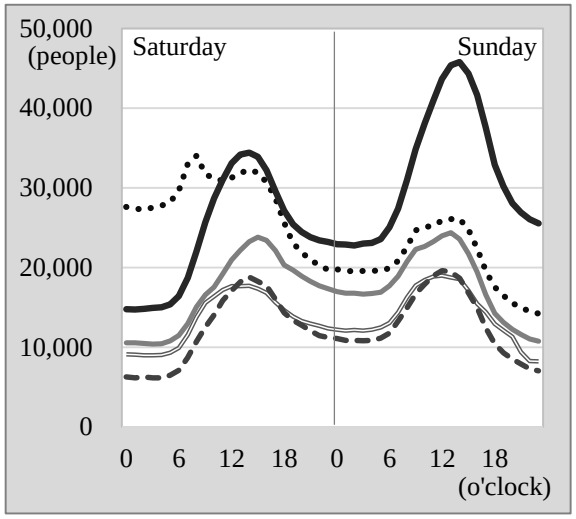

Figure 9. Population distribution chart **(**Hakone-Town, Kanagawa**)**

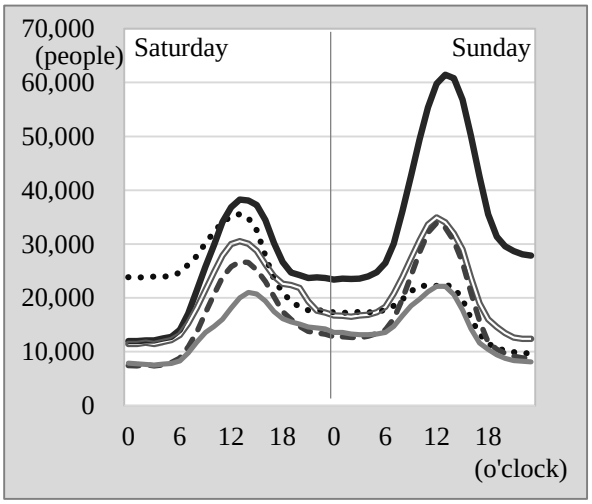

Figure 10. Population distribution chart **(**Nikko-City, Tochigi**)**

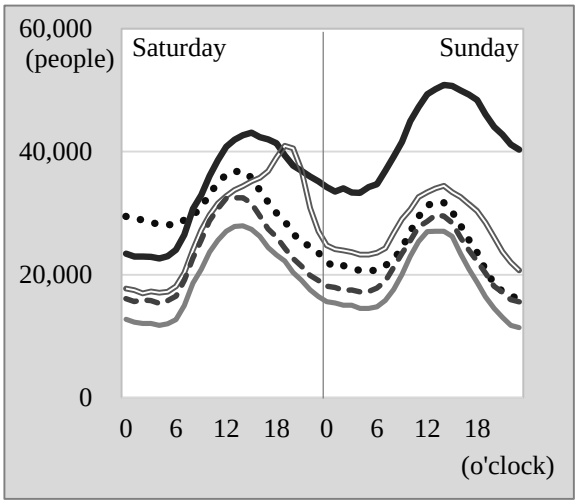

Figure 11. Population distribution chart **(**Hakodate-City, Hokkaido**)**

Himeji-City, Hyogo (Figure 8) is famous for its world heritage, Himeji Castle. The peak is around 13 o'clock and the population at night is small and, there are more population on Saturdays than on Sundays excluding May data.

3.3.3 Cluster C

Cluster C includes many cities for example Hakone, Nikko, and Hakodate, which are famous tourist attractions as examples.

Hakone-Town, Kanagawa (Figure 9) is famous as a hot spring resort and is an area where Mt. Fuji can be seen. The population fluctuates greatly in January and May, the long vacation period. In May, the population tends to increase from Saturday to Sunday because the next day after Sunday is a holiday. On the whole Saturday night is more popular than Sunday night, it is thought that there are many accommodation tourists.

Nikko-City, Tochigi (Figure 10) has the historical architectures including the World-Heritage shrines and temples. According to the figure, the population of long vacation is large, and the population of night time zone on Saturday is large, therefore it is considered that there are many accommodation tourists.

Finally, Hakodate-City, Hokkaido (Figure 11) is shown. A view from Mt. Hakodate is favored by many tourists, and the scenery of the harbor and the townscape of Japanese and Western fashioned houses are distinctive. Compared with the population increase from morning to the peak, the declining population from there is gentle, and it seems that the night population is also relatively large. A rapid population increase was seen on Saturday evenings in August, which is presumed due to the fireworks festival of the Hakodate Port Festival. The effect of the temporary event was confirmed like this.

3.3.4 Summary

Cluster A includes cities located in metropolitan areas, and other areas with large population fluctuations. In these municipalities, the movement of the population during the day is large, but the movement at night is small. Cluster B includes areas famous as regional cities such as Fukuoka, Sapporo and, Sendai. In addition, tourist areas such as Himeji and Izumo are also can be seen. Cluster A and B included day trip type metropolitan cities and regional cities. Cluster C contains several cities famous for hot spring resorts and is considered to have accommodation characteristics. "Daytime-Nighttime Population Ratio" can compare the magnitude of demographics in the daytime, and in the "Nightly Population Ratio", it was possible to classify municipalities as day trip type and accommodation type.

4. RECOMMENDATION FOR TOURISM POLICY

In Chapter 3, characteristics of each tourist areas were shown by confirming the population distribution by time zone and month. In this Chapter, characteristics and problems at these areas are summarized and recommendation for tourism policy is indicated.

First, the peak month of each areas are in May, followed by December and January. The next day is a holiday in May, and the previous day is a holiday in January, so it tends to become a peak month. Therefore, the municipalities where a peak month is May or January is presumed to be tourist areas often visited by people during the consecutive holiday period.

Subsequently, looking at population distribution charts, in areas like cluster A, the population significantly increases during the daytime, but the population at nighttime is small, therefore the population fluctuation is extremely large in a day. In such areas, it is difficult to avoid congestion by shifting time. Among them, tourist areas like Urayasu-City are thought to be influenced by factors such as long holidays, events, weather and so on. It goes without saying that sightseeing is preferred on a day with good conditions, but by analyzing influence factor it is possible to avoid crowded days.

Considering with Ise-City, it is seen that the population of the peak time zone is particularly high. For example, the population of the 13 o'clock of February is 37,679 people, 27,210 people in the 10 o'clock, 24,534 people in the 16 o'clock, there is a large difference between the peak time and the time before and after. Therefore, it is possible to reduce crowded situation by avoiding the peak time. For municipalities that have tourism characteristics like Ise-City, avoiding peak hours will enable tourism with high comfort.

As described above, the population concentration and variability of the tourist areas was confirmed by using "Mobile-Spatial-Statistics". Although this data are aggregated data by area, estimation data of a large population can be utilized, and data conforming to the actual situation can be used. In this research, only the weekend data of the first week of the month is used, but the possibility of utilizing this data in the tourism field is high. For example, if there is data of each year, it is possible to compare the status before and after the project, and if there is daily data, the movement of the population in the tourist areas can be analyzed. Therefore, by analyzing this data together with other data, it is considered to find new possibility of utilization.

5. CONCLUSION

In this research, we used "Mobile-Spatial-Statistics" provided by NTT DOCOMO, Inc. to classify municipalities by making use of the cluster analysis statistics such as "Daytime-Nighttime Population Ratio" and "Nightly population Ratio". Furthermore, we grasped the difference in tourist area characteristics by confirming the population distribution charts by time zone and month in tourist areas for each cluster. The tourist areas had its own features, such as areas with large population fluctuations, many day trippers or accommodation tourists, a high population not only during the day but also at night, affected by weather and events. "Mobile-Spatial-Statistics" are rich in data volume, it is possible to grasp demographics in areas close to the actual situation. It seems that the possibility of its utilization is also high in the tourism field.

REFERENCES

NTT DOCOMO, Inc., Mobile Spatial Statistics Guidelines, Available at: <https://www.nttdocomo.co.jp/english/binary/pdf/service/world/inroaming/inroaming_service/Mobile_Spatial_Statistics_Guidelines.pdf#page=1> [Accessed January 16, 2017]

Furutani, T. (2006) Spatial and Temporal Data Mining of Tourists' Behavior Dynamics by Handheld Positioning Devices -A Case Study in Hakone-. *Journal of the City Planning Institute of Japan,* No.41-3,1-6.

De Cantis, S., Ferrante, M., Kahani, A. and Shoval, N. (2016) Cruise passengers' behavior at the destination: Investigation using GPS technology, *Tourism Management 52*, 133-

150.
Ubukata, Y., Sekimoto, Y. and Horanont, T. (2013) Availability as Tourism Statistical data of Large Scale and Long Term Human Mobility Tracks by GPS: A Study of Ishikawa Pref., *Journal of Japan Society of Civil Engineers, Ser.D3*, I_345-I_352.
Katagiri, U., Shimizu, T. and Kato, S. (2015) Relationship Between the Activity Areas of Foreign Tourists and Evacuation Sites in Tokyo, *Sociotechnica*, Vol.12, 61-70.
Japan Tourism Agency (2015) Tourism Promotion Utilizing Big Data / Survey Analysis of Tourism Behavior Using GPS,
Available at: <http://www.mlit.go.jp/kankocho/shisaku/kankochi/gps.html> [Accessed January 16, 2017]