

ACTIVITY FEATURE EXTRACTION DEVICE AND ACTIVITY FEATURE EXTRACTION METHOD

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to an activity feature extraction device and an activity feature extraction method for analyzing activity route data including movement trajectory information.

2. Description of Related Art

In a recent year, with spread of e-commerce and introduction of remote work, people's activities are becoming less restricted by physical location and time, and are becoming more diverse. In order for a transportation company and a commercial facility operator to implement effective policies amid these social changes, it is important for them to analyze people's actual activities and present them in a format that is understandable to a policymaker.

As a related technique of the present technique, there is Patent Document 1 (JP2016-177452A). The related technique describes a movement route data anonymization device including route data acquisition means for acquiring movement route data including a movement route and a time period during which the movement is made, route

anonymization means for classifying a plurality pieces of the movement route data based on a similarity of the movement routes so that the number of movement route data within the same class is equal to or greater than a first threshold, and unifying the movement routes of the movement route data within the same class into a representative movement route, time anonymization means for abstracting a time period during which the movement is made in the movement route data so that the movements made during the same time period for the movement route data within the same class are equal to or greater than a second threshold, and output means for outputting the movement route data after processing by the route anonymization means and the time anonymization means.

In the aforementioned Patent Document 1 (JP2016-177452A), in anonymizing movement route data, the movement route data including the movement route (location information) and the time period of movement is classified from a perspective of location, and then abstracted in a format that can be effectively used from a perspective of time period.

In case of classifying movement routes from the perspective of location and time period, when classifying movement route data in accordance with actual activity, the appropriate number of classification clusters and the

threshold for time period width for classification from the perspective of time period will change depending on a tendency of movement route data classified into the same cluster from the perspective of location. Similarly, the appropriate number of classification clusters and the threshold for location width for classification from the perspective of location will change depending on a tendency of movement route data classified into the same cluster from the perspective of time period. However, when movement routes are classified into two stages, from the perspective of location and the perspective of time period, a dependency between the location and the time period is not taken into account.

For this reason, when extracting activity features, it is necessary to take into account the location information and time information of the movement trajectory in an integrated manner in order to classify movement routes in accordance with the actual times and places where people are active.

SUMMARY OF THE INVENTION

A representative example of the invention disclosed in this application is as follows. That is, the representative example is an activity feature extraction device that analyzes activity route data in which person

movement trajectory information and time information are associated with each other, where the activity feature extraction device is configured with a computer including an arithmetic device that executes a predetermined arithmetic process and a storage device that the arithmetic device can access, the arithmetic device includes a spatial divergence calculation unit that calculates a spatial divergence representing a spatial difference between the activity route data, the arithmetic device includes a temporal divergence calculation unit that calculates a temporal divergence that represents a temporal difference between the activity route data, the arithmetic device includes a spatiotemporal clustering unit that calculates an activity feature similarity from the spatial divergence and the temporal divergence, and clusters the activity route data using the calculated activity feature similarity, and the arithmetic device includes an output generation unit that outputs information of the cluster generated by the spatiotemporal clustering unit.

According to one aspect of the invention, it is possible to extract activity groups with features that correspond to the actual times and places where people are active, and a policymaker can appropriately consider the type, implementation time, and implementation place of an implemented policy. Issues, configurations, and effects

other than those described above will be made clear through description of an embodiment below.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a functional block diagram illustrating a configuration of an activity feature extraction device according to a first embodiment of the present invention;

Fig. 2 is a block diagram illustrating a hardware configuration of the activity feature extraction device according to the first embodiment of the invention;

Fig. 3 is a diagram illustrating an example of a configuration and data of an activity route database according to the first embodiment of the invention;

Fig. 4 is a flowchart of processing executed by the activity feature extraction device according to the first embodiment of the invention;

Fig. 5 is a flowchart of spatial divergence calculation processing executed by a spatial divergence calculation unit according to the first embodiment of the invention;

Fig. 6 is a flowchart of temporal divergence calculation processing executed by a temporal divergence calculation unit according to the first embodiment of the invention;

Fig. 7 is a flowchart of spatiotemporal clustering

processing executed by a spatiotemporal clustering unit according to the first embodiment of the invention;

Fig. 8 is a flowchart of feature extraction processing executed by a feature extraction unit according to the first embodiment of the invention;

Fig. 9 is a diagram illustrating an example of a setting input screen output from an output device of the activity feature extraction device according to the first embodiment of the invention;

Fig. 10 is a diagram illustrating an example of an activity feature display screen output from the output device of the activity feature extraction device according to the first embodiment of the invention;

Fig. 11 is a functional block diagram illustrating a configuration of an activity feature extraction device of a second embodiment according to the invention;

Fig. 12 is a flowchart of processing executed by the activity feature extraction device of the second embodiment according to the invention;

Fig. 13 is a diagram illustrating an example of a person attribute input screen output from an output device of the activity feature extraction device according to the second embodiment of the invention; and

Fig. 14 is a diagram illustrating an example of a spot recommendation screen output from the output device of the

activity feature extraction device according to the second embodiment of the invention.

DESCRIPTION OF EMBODIMENTS

First Embodiment

Hereinafter, an embodiment of the present invention will be described with reference to the drawings.

In this embodiment, an activity feature extraction device 100 clusters activity route data, including a person's movement trajectory information and time information associated with the movement trajectory information, using an activity feature similarity based on both spatial difference and temporal difference between the activity route data, to extract activity groups having characteristics that correspond to the actual times and places where people are active. For this reason, this embodiment can support a policymaker in appropriately considering the type, implementation time, and implementation location of an implemented policy.

Fig. 1 is a functional block diagram illustrating a configuration of the activity feature extraction device 100 according to a first embodiment of the invention.

The activity feature extraction device 100 of this embodiment includes an input device 101, an output device 102, and a server computer 103.

The input device 101 is an input interface such as a mouse, keyboard, or touch device for transmitting a user operation to the server computer 103. The output device 102 is an output interface configured with a display device such as a liquid crystal display, and is an output interface for displaying data output by the server computer 103 and for an interactive operation with the user. For example, the input device 101 and the output device 102 may be integrated by using a so-called touch panel or the like. The output device 102 displays a setting input screen 901 (Fig. 9), an activity feature display screen 1001 (Fig. 10), a person feature input screen (Fig. 13), and a spot recommendation screen 1401 (Fig. 14), which will be described below. The input device 101 accepts a user operation on input fields and buttons on each screen. The output device 102 may be the display device described above, a printer that prints processing results on paper, or a data output interface that transfers processing results to another system. Also, as in the second embodiment, the input device 101 and the output device 102 may be terminals connected to the activity feature extraction device 100 via a network 113. In this case, the activity feature extraction device 100 may have a web server function, and the terminals may access the activity feature extraction device 100 using a predetermined protocol (for example,

http).

The server computer 103 has, as its functional configuration, a setting input unit 104, an activity route data acquisition unit 105, a spatial divergence calculation unit 106, a temporal divergence calculation unit 107, a spatiotemporal clustering unit 108, a feature extraction unit 109, an output generation unit 110, and an activity route database 111.

The setting input unit 104 acquires target area information, target date and time information, and classification setting information input to the input device 101, sends the target area information and target date and time information to the activity route data acquisition unit 105, and transmits the classification setting information input to the input device 101 to the spatiotemporal clustering unit 108. The target area information may include at least one of an address of the target area, a distance between a facility of interest and the facility, and a coordinate group representing a polygon shape of a target area range. The target date and time information may include at least one of the date, day of the week, temperature conditions, seasonal conditions, weather conditions, and day of the week conditions. The classification setting information may include at least one of the number of clusters, inter-data similarity threshold,

minimum number of data, model structure, and model parameters.

The activity route data acquisition unit 105 acquires the target area information and the target date and time information from the setting input unit 104, extracts an activity route data group recorded within an area specified by the target area information and within a date and time specified by the target date and time information from the activity route data stored in an activity route database 114, and transmits the data group to the spatial divergence calculation unit 106 and the temporal divergence calculation unit 107, respectively. Here, the activity route data includes at least person movement trajectory information and time information associated with the movement trajectory information.

The spatial divergence calculation unit 106 acquires the activity route data group from the activity route data acquisition unit 105, and calculates spatial divergence for all combinations of two activity route data included in the activity route data group. Here, the spatial divergence represents a geographical difference in movement trajectories between the two activity route data.

The spatial divergence is calculated, for example, as in Dynamic Time Warping (DTW), by searching for a closest pair group which is observation point pair that can be

determined to be closest between all observation points of a certain movement trajectory and all observation points of a movement trajectory to be compared while satisfying a constraint condition of an observation order, and, for example, between activity route data A and B, as illustrated in Equation (1), using a distance between an observation point $p_{A(k)}$ included in the activity route data A and an observation point $p_{B(k)}$ included in the activity route data B corresponding to a closest pair k, and a geographical condition coefficient a_k determined by features of a location where both observation points are located, and using a measure that can take into account a difference between movement trajectories of activity route data while taking into account an influence of observation errors that depend on a surrounding environment such as the size of the facility where each observation point is located, transportation means, indoor/outdoor, and radio wave environment.

Thereafter, the spatial divergence calculation unit 106 sends the closest pair group to the temporal divergence calculation unit 107, and transmits the calculated spatial divergence to the spatiotemporal clustering unit 108.

[Equation 1]

$$D_L(\mathbf{A}, \mathbf{B}) = \sum_k a_k \|p_{A(k)}, p_{B(k)}\| \dots (1)$$

The temporal divergence calculation unit 107 acquires

the activity route data group from the activity route data acquisition unit 105, acquires the closest pair group from the spatial divergence calculation unit 106, and calculates the temporal divergence for all combinations of two activity route data included in the acquired activity route data group. Here, the temporal divergence represents a time difference between two activity route data.

The temporal divergence is calculated, for example, between activity route data A and B, using a measure that can take into account a difference between activity route data in at least one of the following aspects: a time period during which the activity occurred, transportation means at the time, and whether the activity is passed through or stayed, using a speed $v_{A(k)}$ of the observation point $p_{A(k)}$ and a speed $v_{B(k)}$ of the observation point $p_{B(k)}$ corresponding to a closest pair k , and an observation time difference Δt_k between both observation points, as illustrated in Equation (2).

The temporal divergence calculation unit 107 then transmits the calculated temporal divergence to the spatiotemporal clustering unit 108.

[Equation 2]

$$D_T(\mathbf{A}, \mathbf{B}) = \sum_k \Delta t_k \frac{v_{A(k)} + v_{B(k)}}{2} \dots (2)$$

The spatiotemporal clustering unit 108 acquires the

classification setting information from the setting input unit 104, acquires the spatial divergence from the spatial divergence calculation unit 106, acquires temporal divergence from the temporal divergence calculation unit 107, and clusters the activity route data group using activity feature similarity based on the spatial divergence and temporal divergence. The activity feature similarity may be dissimilarity, and the temporal divergence calculation unit 107 converts the temporal divergence into a distance dimension by defining it in a form in which the observed time difference is multiplied by the speed. Thus, for example, as in Equation (3), the activity feature similarity can be calculated by directly adding the spatial divergence $D_L(A, B)$ and the temporal divergence $D_T(A, B)$ without considering the difference in dimension.

[Equation 3]

$$D(A, B) = D_L(A, B) + D_T(A, B) \dots (3)$$

A calculation method of clustering may be changed according to the acquired classification setting information. When the number of clusters is acquired, for example, k-means is used to classify the activity route data into the number of clusters based on the activity feature similarity, and an activity route cluster is generated. On the other hand, when an inter-data similarity threshold is acquired, for example, a density-based clustering method

such as Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is used to cluster the activity route data based on the activity feature similarity, and an activity route cluster is generated so that the activity feature similarity between the activity route data in each cluster satisfies a condition of the inter-data similarity threshold. Then, the activity route cluster is transmitted to the feature extraction unit 109.

The feature extraction unit 109 acquires the activity route clusters from the spatiotemporal clustering unit 108 and extracts cluster feature information from the activity route data in each activity route cluster. Here, the cluster feature information is information on tendencies of people who generated the activity route data in each activity route cluster, and preferably includes at least one of age, gender, occupation, starting time of work, finishing time of work, and hobbies.

When the activity route data is provided with person attribute information, the cluster feature information may be extracted by aggregating or existing statistical processing the person attribute information of the person who generated the activity route data in the activity route cluster. When the activity route data is not provided with person attribute information, for example, map information including facility information and transportation network

information of the planning target area is acquired from the outside via the network 113 and a network interface device 112, and the facilities through which the activity route data passed and stayed and the transportation means estimated to be used are extracted, and a proportion of people who take the activity route data may be extracted by aggregating or existing statistical processing from user attributes of the facility and/or the transportation means and time information of the activity route data. Then, the feature extraction unit 109 transmits the activity route cluster and the cluster feature information to the output generation unit 110.

The output generation unit 110 acquires the activity route clusters and the cluster feature information from the feature extraction unit 109, creates the activity feature display screen 1001 (Fig. 10), and outputs it to the output device 102. The configuration of the activity feature display screen 1001 will be described below with reference to Fig. 10.

The activity route database 111 is a database for storing activity route data. The activity route data acquisition unit 105 accesses the activity route database 111 when executing an extraction process. The configuration of the activity route database 111 will be described below with reference to Fig. 3.

Fig. 2 is a block diagram illustrating a hardware configuration of the activity feature extraction device 100 according to the first embodiment of the invention.

The server computer 103 is, for example, a general-purpose computer having a processor 201 and a storage device 202 connected to each other. The storage device 202 is composed of any type of storage medium. For example, the storage device 202 may include a semiconductor memory and a hard disk drive.

In the illustrated example, functional units such as the setting input unit 104, activity route data acquisition unit 105, spatial divergence calculation unit 106, temporal divergence calculation unit 107, spatiotemporal clustering unit 108, and feature extraction unit 109 illustrated in Fig. 1 are realized by the processor 201 executing a processing program 203 stored in the storage device 202.

In other words, in this embodiment, the processing executed by each of the functional units described above is actually executed by the processor 201 following instructions written in the processing program 203. Furthermore, the activity route database 111 is configured in the storage device 202. Furthermore, the display by the output device 102 is executed by the processor 201 generating display data for display and outputting it to the output device 102, and the output device 102 performing

display according to the display data.

The server computer 103 further includes the network interface device (NIF) 112 connected to the processor 201. The network interface device 112 is a communication interface that controls communication with other devices according to a predetermined protocol.

The program executed by the processor 201 is provided to the activity feature extraction device 100 via removable media (CD-ROM, flash memory, or the like) or the network 113, and is stored in the storage device 202 having a non-volatile characteristic, which is a non-transitory storage medium. For this reason, the server computer 103 may have an interface for reading data from the removable media.

The activity feature extraction device 100 is a computer system configured on one physical computer, or on a plurality of logically or physically configured computers, and may operate on a virtual computer constructed on a plurality of physical computer resources. For example, each functional unit may operate on a separate physical or logical computer, or a plurality of functional units may be combined to operate on a single physical or logical computer.

Fig. 3 is a diagram illustrating an example of the configuration and data of the activity route database 111 in the first embodiment of the invention. Although an example in which the activity route database 111 is

configured in a table format is illustrated, the data format is not limited to the table format and may be another format.

The activity route database 111 includes activity route data management information 300 illustrated in Fig. 3. The table configuration and field configuration of each table illustrated in Fig. 3 are preferred configurations for the invention, but tables and fields may be added depending on the application.

The activity route data management information 300 includes an agent ID field 301, a data acquisition time field 302, a latitude field 303, a longitude field 304, and a person attribute field 305.

The agent ID field 301 holds identification information (hereinafter referred to as a person ID) of the person who performed the action of each piece of trajectory data.

The data acquisition time field 302 holds information about the time when a trajectory point is acquired.

The latitude field 303 holds the latitude information of a position where the trajectory point is acquired.

The longitude field 304 holds the longitude information of a position where the trajectory point is acquired.

The person attribute field 305 holds person attribute information of a user corresponding to the person ID

described above. Here, the person attribute information includes age, gender, occupation, start and finish times of work, hobbies, and the like, and available information may be expressed as an array of which elements are numerical values or one-hot vectors.

Fig. 4 is a flowchart of processing executed by the activity feature extraction device 100 according to the first embodiment of the invention.

First, when a user inputs target area information, target date and time information, and classification setting information into the input device 101, the input device 101 transmits the input information to the setting input unit 104 (step S401).

Next, the setting input unit 104 transmits the target area information and the target date and time information to the activity route data acquisition unit 105 (step S402), and transmits the classification setting information to the spatiotemporal clustering unit 108 (step S403).

Next, the activity route data acquisition unit 105 extracts an activity route data group from the activity route database 111 according to conditions of the target area information and target date and time information, and transmits the target activity route data group to the spatial divergence calculation unit 106 and the temporal divergence calculation unit 107 (step S404).

Next, the spatial divergence calculation unit 106 extracts the closest pair group between the activity route data groups, calculates the spatial divergence, and then transmits the closest pair group to the temporal divergence calculation unit 107 (step S405), and transmits the calculated spatial divergence to the spatiotemporal clustering unit 108 (step S406). Details of a spatial divergence calculation processing will be described below with reference to Fig. 5.

Next, the temporal divergence calculation unit 107 calculates the temporal divergence using the closest pair group and transmits the calculated temporal divergence to the spatiotemporal clustering unit 108 (step S407). Details of a temporal divergence calculation processing will be described below with reference to Fig. 6.

Next, the spatiotemporal clustering unit 108 calculates the activity feature similarity from the spatial divergence and the temporal divergence, generates an activity route cluster according to the activity feature similarity, and then transmits the activity route cluster to the feature extraction unit 109 (step S408). Details of a spatiotemporal clustering processing will be described below with reference to Fig. 7.

Next, the feature extraction unit 109 extracts cluster feature information from the activity route cluster and

transmits the activity route cluster and the cluster feature information to the output generation unit 110 (step S409). Details of a feature extraction processing will be described below with reference to Fig. 8.

Finally, the output generation unit 110 creates the activity feature display screen 1001 (Fig. 10) and transmits it to the output device 102 (step S410). The user can check processing results by looking at the activity feature display screen 1001 displayed on the output device 102.

Fig. 5 is a flowchart of the spatial divergence calculation processing executed by the spatial divergence calculation unit 106 according to the first embodiment of the invention.

First, the spatial divergence calculation unit 106 acquires an activity route data group from the activity route data acquisition unit 105 (step S501).

Next, the spatial divergence calculation unit 106 repeatedly executes processes of steps S503 to S506 for all combinations of two activity route data A and B in the acquired activity route data group that do not include overlaps (steps S502, S507).

In the repetition, first, the spatial divergence calculation unit 106 calculates a distance between the observation points using the latitude and longitude information between each observation point in the activity

route data A and B (step S503). Next, the spatial divergence calculation unit 106 calculates a cumulative distance matrix (step S504) and searches for an optimal route that minimizes a cumulative distance (step S505). For example, dynamic programming can be used to calculate the cumulative distance matrix.

Then, the spatial divergence calculation unit 106 extracts the closest pair group of observation points between the activity route data A and B from the optimal route, and sets the cumulative distance of the optimal route as the spatial divergence (S506). Then, the processing moves to the next activity route data pair (S507).

Finally, the spatial divergence calculation unit 106 outputs all the closest pair groups and their spatial divergences (step S508).

Fig. 6 is a flowchart of the temporal divergence calculation processing executed by the temporal divergence calculation unit 107 according to the first embodiment of the invention.

First, the temporal divergence calculation unit 107 acquires the closest pair group of each activity route data from the spatial divergence calculation unit 106 (step S601).

Next, the temporal divergence calculation unit 107 repeatedly executes processes of steps S603 to S605 for all combinations of the two activity route data A and B in the

acquired activity route data group that do not include overlaps (steps S602 and S606).

In the repetition, first, the temporal divergence calculation unit 107 calculates an observed time difference between the closest pair of the activity route data A and B (step S603), and multiplies the observed time difference by an average of two points of the speed at both observation points of the closest pair to calculate the temporal divergence between the closest pair (step S604). Then, the temporal divergence calculation unit 107 calculates the temporal divergence between all closest pairs, and calculates the temporal divergence between the activity route data A and B by adding up the calculated temporal divergences (step S605). Finally, the temporal divergence calculation unit 107 outputs all the calculated temporal divergences (step S607).

Fig. 7 is a flowchart of the spatiotemporal clustering processing executed by the spatiotemporal clustering unit 108 in the first embodiment of the invention, and illustrates processing when an inter-data similarity threshold and the minimum number of data are given.

First, the spatiotemporal clustering unit 108 acquires the spatial divergence from the spatial divergence calculation unit 106 and the temporal divergence from the temporal divergence calculation unit 107 (step S701).

Next, the spatiotemporal clustering unit 108 adds the spatial divergence and the temporal divergence for each pair of the activity route data to calculate the activity feature similarity between the activity route data (step S702). Since the temporal divergence is converted into a dimension of spatial divergence using the speed, a simple addition operation can be used to calculate the activity feature similarity that includes the features of both spatial divergence and temporal divergence. The calculated activity feature similarity then enables clustering that takes into account the duration of stay and the means of transportation.

Then, the spatiotemporal clustering unit 108 randomly selects one activity route data of interest as an initial value from the activity route data group (step S703), and repeatedly executes processes of steps S705 to S710 until all activity route data are assigned to clusters or noise (steps S704, S711).

In the repetition, first, the spatiotemporal clustering unit 108 extracts all activity route data of which the activity feature similarity with the activity route data of interest is equal to or greater than the inter-data similarity threshold (step S705). In this case, when the number of relevant activity route data is equal to or greater than the minimum number of data (step S706: Yes), a new cluster is assigned to the activity route data of

interest and the relevant activity route data (step S707), and the same processing as in steps S705 and S707 is performed on all relevant activity route data, and this is repeated until there is no unprocessed relevant activity route data.

When the number of relevant activity route data is smaller than the minimum number of data (step S706: No), the spatiotemporal clustering unit 108 classifies the activity route data of interest into noise that does not form a cluster.

After the process of step S708 or S709 is completed, when there is activity route data not assigned to any cluster, the spatiotemporal clustering unit 108 randomly selects one activity route data from among them as activity route data of interest.

Finally, the spatiotemporal clustering unit 108 outputs all the activity route clusters (step S712).

Fig. 8 is a flowchart of the feature extraction processing executed by the feature extraction unit 109 according to the first embodiment of the invention.

First, the feature extraction unit 109 acquires the activity route clusters from the spatiotemporal clustering unit 108 (step S801). In this case, when a person attribute is assigned to the activity route data included in each activity route cluster (step S802: Yes), the process of step

S804 is repeatedly executed for all acquired activity route clusters (steps S803, S805).

In the repetition, the feature extraction unit 109 extracts cluster features representative of the activity route cluster by aggregating the person attributes in the activity route cluster or by performing statistical processing using an existing statistical method (step S804).

When a person attribute is not assigned to the activity route data included in each activity route cluster (step S802: No), the feature extraction unit 109 executes the processes of steps S807 to S811 for all acquired activity route clusters (steps S806, S812). Then, the feature extraction unit 109 executes the processes of steps S808 to S809 for the activity route data included in the activity route cluster (steps S807, S810).

In the repetition, first, the feature extraction unit 109 acquires map information including facility information and transportation network information of the planning target area from the outside via the network 113 and the network interface device 112, and extracts facilities passed or stayed in by the activity route data and transportation means used in the activity route data (step S808). For example, a threshold for a distance from the facility used to determine whether data passed or stayed is set, and the movement trajectory information and time information of the

activity route data are compared with the set threshold to extract the facilities passed or stayed in by the activity route data. In addition, for example, an existing transportation mode estimation method is used to extract the transportation means used in the activity route data from the movement trajectory information and time information of the activity route data.

Next, the feature extraction unit 109 acquires user attribute information for the extracted relevant facility and relevant transportation means at the relevant time (step S809).

Then, the feature extraction unit 109 extracts user attribute information for all activity route data, and then extracts cluster features that represent the activity route cluster by aggregating the extracted user attribute information or by performing statistical processing using existing statistical methods (step S811).

Finally, the feature extraction unit 109 outputs all the activity route clusters and the cluster feature information (step S813).

Fig. 9 is a diagram illustrating an example of the setting input screen 901 output from the output device 102 of the activity feature extraction device 100 according to the first embodiment of the invention.

The setting input screen 901 illustrated in Fig. 9

includes a target area input area 902, a target area information input area 903, a target date and time information input area 904, a classification setting information input area 905, and a map display management area 906.

In the target area input area 902, a target area marker 902A and an orientation display 902B are displayed on a map.

When the user specifies a range to be analyzed on the map in the target area input area 902, or inputs target area information into the target area information input area 903, the target area marker 902A surrounding a periphery of the specified area is displayed. Then, planning target site information is transmitted from the input device 101 to the setting input unit 104.

The user can also change the scale and display position by specifying on the target area input area 902 using a method different from that used when selecting the target area. In this case, it is preferable that the user can change the angle of the displayed map by specifying on the target area input area 902 using an event to which a key operation, mouse click, or touch method different from that described above is assigned. The orientation display 902B indicates the orientation of the map displayed in the target area input area 902.

By displaying the target area input area 902 in this

way, characteristic locations such as major facilities and rivers can be displayed together with the analysis target area, making it easy to find and select the analysis target area on the map.

The target area information input area 903 accepts target area information input by the user. The target area may be specified by only address information input in the target area information input area 903, or may be specified by a facility of interest selected in the target area input area 902 and an area radius (distance from the facility) input in the target area information input area 903.

The target date and time information input area 904 includes a date and time setting area 904A and a condition setting area 904B.

The user inputs a period or time in the date and time setting area 904A as the date and time to be analyzed. In the condition setting area 904B, one or more of temperature, season, weather, and day of the week can be selected and input, and target date and time information that meets the specified condition within the period specified in the date and time setting area 904A is registered in the setting input unit 104. When no period is input in the date and time setting area 904A, activity route data that meets the condition input in the condition setting area 904B from among the activity route data for the entire period stored

in the activity route database 114 will be the analysis target. When no condition is input in the condition setting area 904B, all activity route data within the period specified in the date and time setting area 904A will be the analysis target.

In the classification setting information input area 905, selection of a classification method to be used for activity feature extraction and classification setting information required for the selected classification method are input. When a classification method is selected, only the setting information required for the selected method is accepted as input.

After all of the input of the target area information, target date and time information, and classification setting information is completed, an extraction start button 905A is activated, and the processing of the setting input unit 104 is executed by operating the extraction start button 905A.

In the map display management area 906, a scale reduction button 906A, a scale increase button 906B, a scale 906C, and a home position transition button 906D are displayed.

In the map display management area 906, the user can operate the scale reduction button 906A with the cursor or touch to reduce the scale of the map displayed in the target

area input area 902 and display a wider area. Also, the user can operate the scale increase button 906B to increase the scale of the map displayed in the target area input area 902 and display a narrower area in detail. The scale 906C displays a distance on the map displayed in the target area input area 902. Also, the user can operate the home position transition button 906D with the cursor or touch to display the map with the position, scale, and orientation displayed in an initial screen in the target area input area 902.

Fig. 10 is a diagram illustrating an example of the activity feature display screen 1001 output from the output device 102 of the activity feature extraction device 100 according to the first embodiment of the invention.

The activity feature display screen 1001 illustrated in Fig. 10 includes a cluster display area 1002, a cluster feature display area 1003, and an animation control area 1004.

In the cluster display area 1002, activity route data 1002A for a target area and a target date and time, and a cluster 1002B to which the activity route data belongs are displayed on a map displayed in the cluster display area 1002.

Here, time information accompanying the activity route data may be displayed in a display format according to the activity route data, such as by setting a time axis

perpendicular to a map surface, plotting each observation point of the activity route data on a three-dimensional coordinate system of latitude, longitude, and time, and displaying a trajectory connecting the observation points, or by displaying the activity route data corresponding to the time displayed in the animation control area 1004 through animation.

Furthermore, the user can change the scale, display position, and viewpoint for displaying three-dimensional coordinates in the cluster display area 1002 by operating the cursor or touching. In this case, it is preferable that the user can change the angle of the map displayed by operating on the cluster display area 1002 using an event to which a key operation, mouse click, or touch method different from those described above is assigned.

Furthermore, the cluster 1002B displays an oval surrounding the activity route data belonging to each cluster, but the activity route data may be displayed in a distinguishable manner using a display format (different colors, shades of color, or the like) according to the cluster.

The cluster feature display area 1003 appears when a cluster is selected by the cursor or touch, and displays the cluster feature information extracted by the feature extraction unit 109. The cluster feature information allows

a policymaker to easily understand people's characteristics and appropriately narrow down marketing target persons.

The animation control area 1004 includes a play button 1004A that controls playing and stopping of the animation to control the display of the activity route data 1002A and the cluster 1002B to which the activity route data belongs, a seek bar 1004B that indicates the display timing and allows the display timing to be specified, a reverse play button 1004C that allows the animation to be rewound, and a time zone specification field 1004D in which the display time zone can be input.

The animation control area 1004 allows movement trajectories with high spatial similarity to be displayed as belonging to the same cluster even when the activity times are different, and the cluster features can be confirmed (for example, cluster A). Furthermore, the animation control area 1004 allows movement trajectories with high temporal similarity to be displayed as belonging to the same cluster even when the movement trajectories are different, and the cluster features can be confirmed (for example, cluster E).

In the embodiment described above, the temporal divergence is multiplied by the speed to calculate the temporal divergence having the same dimension as the spatial divergence, but the spatial divergence may be divided by

the speed to calculate the spatial divergence having the same dimension as the temporal divergence.

<Operations and Effects of the First Embodiment>

As described above, the activity feature extraction device 100 of this embodiment can extract activity groups having features that correspond to the actual times and places where people are active, taking into account differences in facilities used and transportation means, so that the policymaker can appropriately consider the type, implementation time, and implementation location of the policy to be implemented using the features of the activity groups.

In addition, the activity feature extraction device 100 of this embodiment can provide the policymaker with an activity status of the area under consideration, which is the basis for determining the type, implementation time, and implementation location of the policy research, from a consideration stage or early stage of the policy when there is no data showing the effect of the implementation of the policy. This reduces a possibility that the effect of the policy will be less than expected after implementation, and reduces cost of modifying the policy.

(Second Embodiment)

An activity feature extraction device 100 according to a second embodiment will be described. In this

embodiment, the same process as in the first embodiment is performed up to the extraction of cluster feature information, and further, the user inputs person attribute information of the user himself or herself or a related person, and the cluster to which the user himself or herself belongs and/or the cluster to which the related person belongs are calculated using the cluster feature information, and facility information and transportation information are recommended taking into account the activity features of the user himself or herself and/or the related person. In the second embodiment, differences from the first embodiment will be mainly described, and the same configurations and processes as in the first embodiment will be assigned the same reference numerals, and their description will be omitted.

The second embodiment is basically the same as the first embodiment, but has a configuration added to the activity feature extraction device 100 of the first embodiment illustrated in Fig. 1, and further differs in the processing executed by the activity feature extraction device 100 illustrated in Fig. 4 and the display screen created by the output generation unit 110. Other configurations and processing are the same as those of the first embodiment, so descriptions thereof will be omitted. The configuration added to the activity feature extraction

device 100 of the first embodiment will be described with reference to Fig. 11, the flowchart of the processing of the second embodiment will be described with reference to Fig. 12, and the screen created by the output generation unit 110 will be described with reference to Fig. 13 and Fig. 14.

Fig. 11 is a functional block diagram illustrating a configuration of the activity feature extraction device 100 according to the second embodiment of the invention.

The server computer 103 includes a person attribute input unit 1101, a belonging cluster calculation unit 1102, and a recommended spot extraction unit 1103 in addition to the configurations in the first embodiment.

The person attribute input unit 1101 acquires person attribute information input by the user from the input device 101 and transmits it to the belonging cluster calculation unit 1102. Here, the person attribute information has at least one of the activity route data and person attribute information of a target person. Here, the target person may include not only the user himself or herself but also a person related to the user. When the person attribute information of the target person has already been input, for example, a person ID of the target person may be input and approval for the use of the input person attribute information may be obtained from the target

person.

The feature extraction unit 109 acquires activity route clusters from the spatiotemporal clustering unit 108, extracts cluster feature information from activity route data in each activity route cluster, and transmits the activity route clusters and the cluster feature information to the output generation unit 110 and the belonging cluster calculation unit 1102.

The belonging cluster calculation unit 1102 acquires the activity route clusters and the cluster feature information from the feature extraction unit 109, acquires person attribute information from the person attribute input unit 1101, calculates a belonging probability of a person having the acquired person attribute information to each activity route cluster, assigns the person to the activity route cluster with the highest probability, and determines the belonging cluster information.

The belonging probability to each activity route cluster can be calculated by a following method. For example, when the person attribute information includes activity route data, the activity feature similarity with the activity route data included in each activity route cluster is calculated using the same method as the spatiotemporal clustering unit 108, and the belonging probability is calculated to be proportional to an average

activity feature similarity with the activity route data in each activity route cluster. Also, when the person attribute information includes only person attribute information (no activity route data is included), for example, the cluster feature information of each activity cluster is compared with the person attribute information, and the belonging probability is calculated to be proportional to the attribute similarity between the cluster feature information and the person attribute information. Here, the attribute similarity may be calculated by, for example, converting each item of person's tendency information included in the cluster feature information and the person attribute information into a one-hot vector and calculating a cosine similarity between them. Then, the belonging cluster information is transmitted to the recommended spot extraction unit 1103.

The recommended spot extraction unit 1103 acquires the belonging cluster information from the belonging cluster calculation unit 1102, and calculates recommended spots according to the target person's person attributes.

Recommended spots can be calculated using the following method. For example, all facility information visited by the activity route data included in the cluster to which the target person is determined to belong is weighted by the number of visits to the facility, a matching

rate with the target person's preferences is calculated, and recommended spot information is extracted according to the calculated matching rate.

In this case, when the target persons include one or more related persons in addition to the user himself or herself, similarly for the related person, all facility information visited by the activity route data included in the cluster to which the related person is determined to belong is weighted by the number of visits to the facility, the matching rate for the related person's preferences is calculated, the matching rate is multiplied for each facility to obtain the final matching rate, and recommended spot information is calculated according to the matching rates. The recommended spot information is then transmitted to the output generation unit 110.

The output generation unit 110 of the second embodiment acquires the recommended spot information from the recommended spot extraction unit 1103, creates the spot recommendation screen 1401, and outputs it to the output device 102. The configuration of the spot recommendation screen 1401 will be described below with reference to Fig. 14.

Fig. 12 is a flowchart of processing executed by the activity feature extraction device 100 according to the second embodiment of the invention.

First, when a user inputs person attribute information of the user himself or herself or a related person to the input device 101 (person attribute input screen 1301 (Fig. 13)), the input device 101 transmits the input person attribute information to the person attribute input unit 1101 (step S1201).

Next, the person attribute input unit 1101 transmits the input person attribute information to the belonging cluster calculation unit 1102 (step S1202).

The feature extraction unit 109 also transmits the activity route cluster and the cluster feature information to the belonging cluster calculation unit 1102 (step S1203). Here, the steps up to the calculation of the activity route cluster and the cluster feature information in the first embodiment (steps S401 to S409) are executed before step S1203.

Next, the belonging cluster calculation unit 1102 calculates the cluster to which the target person of the person attribute information belongs from the cluster feature information of each activity route cluster and the person attribute information, and transmits it to the recommended spot extraction unit 1103 (step S1204).

Next, the recommended spot extraction unit 1103 extracts recommended spots and transmits them to the output generation unit 110 (step S1205).

Finally, the output generation unit 110 creates the spot recommendation screen 1401 (Fig. 14) and transmits it to the output device 102 (step S1206). The user can check processing results by viewing the spot recommendation screen 1401 displayed on the output device 102.

Fig. 13 is a diagram illustrating an example of the person attribute input screen 1301 output from the output device 102 of the activity feature extraction device 100 according to the second embodiment of the invention.

The person attribute input screen 1301 illustrated in Fig. 13 includes a person attribute information input area 1302 and a related person information input area 1303.

The person attribute information input area 1302 accepts input of the user's own person attributes by the user.

The related person information input area 1303 accepts user input of person attributes of the related person of the user. The input of related person information is optional. When the person attribute information of the target person has already been registered in the activity feature extraction device 100, the user may need to only input a person ID as the identification number of the target person. On the other hand, when the person attribute information of the target person has not already been input in the activity feature extraction device 100, the user

inputs the target person's person attributes in the related person information input area 1303.

Fig. 14 is a diagram illustrating an example of the spot recommendation screen 1401 output from the output device 102 of the activity feature extraction device 100 according to the second embodiment of the invention.

The spot recommendation screen 1401 illustrated in Fig. 14 includes a recommended spot display area 1402 and a recommended spot information display area 1403. A recommended spot pin 1402A is displayed on the recommended spot display area 1402.

The recommended spot pin 1402A shows, on a map a position, of the recommended spot acquired from the recommended spot extraction unit 1103 and a recommendation level according to the matching rate, and may be displayed in a display format suitable for displaying location information, such as a pin. The recommendation level may be a value determined based on the matching rate, and the matching rate itself or a ranking of the matching rates may be displayed.

The recommended spot information display area 1403 displays recommended spot information. In addition to information about the spot itself, the recommended spot information may display the time periods when activity route data included in the activity cluster to which the user and

the related person belong makes a frequent visit, and a level of interest of the related person.

<Operations and Effects of Second Embodiment>

As described above, the activity feature extraction device 100 of this embodiment can extract activity groups that have features that correspond to the actual times and places where people are active, taking into account differences in facilities used and transportation means, and associate the activity groups with the user, so that the user can learn spot information that matches the tendencies of his or her activity times and activity places, including places that the user has never visited. In addition, the user can learn the spot information without providing his or her own location information.

The invention is not limited to the above-described embodiments, and includes various modified examples and equivalent configurations within the spirit of the appended claims. For example, the above-described embodiments are described in detail to clearly describe the invention, and the invention is not necessarily limited to having all of the configurations described. Furthermore, a part of the configuration of one embodiment may be replaced with a configuration of another embodiment. Furthermore, a configuration of another embodiment may be added to the configuration of one embodiment. Furthermore, a part of

the configuration of each embodiment may be added, deleted, or replaced with other configurations.

In addition, each of the above-described configurations, functions, processing units, processing means, and the like may be realized in hardware, for example by designing some or all of them as an integrated circuit, or may be realized in software by a processor interpreting and executing a program that realizes each function.

Information such as programs, tables, files, and the like that realize each function can be stored in a storage device such as a memory, a hard disk, or a Solid State Drive (SSD), or in a recording medium such as an IC card, an SD card, or a DVD.

Also, the control lines and information lines shown are those considered necessary for description, and do not necessarily show all control lines and information lines necessary for implementation. In reality, almost all components are considered to be interconnected.

What is claimed is:

1. An activity feature extraction device that analyzes activity route data in which person movement trajectory information and time information are associated with each other, wherein

the activity feature extraction device is configured with a computer including an arithmetic device that executes a predetermined arithmetic process and a storage device that the arithmetic device can access,

the arithmetic device includes a spatial divergence calculation unit that calculates a spatial divergence representing a spatial difference between the activity route data,

the arithmetic device includes a temporal divergence calculation unit that calculates a temporal divergence representing a temporal difference between the activity route data,

the arithmetic device includes a spatiotemporal clustering unit that calculates an activity feature similarity from the spatial divergence and the temporal divergence, and clusters the activity route data using the calculated activity feature similarity, and

the arithmetic device includes an output generation unit that outputs information of a cluster generated by the spatiotemporal clustering unit.

2. The activity feature extraction device according to claim 1, wherein

the temporal divergence calculation unit calculates the temporal divergence or the spatial divergence by taking into account speed information associated with the movement trajectory information.

3. The activity feature extraction device according to claim 2, wherein

the temporal divergence calculation unit multiplies the temporal divergence by the speed information to calculate a temporal divergence having the same dimension as the spatial divergence, or divides the spatial divergence by the speed information to calculate the spatial divergence having the same dimension as the temporal divergence.

4. The activity feature extraction device according to claim 2 or 3, wherein

the spatial divergence calculation unit determines a corresponding pair of observation points included in the movement trajectory information among the activity route data, and calculates the spatial divergence by sum of spatial differences between the corresponding pairs, and

the temporal divergence calculation unit calculates

the temporal divergence based on a difference in the time information of the corresponding pair and movement speed at the observation points that form the corresponding pair.

5. The activity feature extraction device according to claim 2 or 3, wherein

the spatial divergence calculation unit calculates the spatial divergence using features of points where people pass or stay in the movement trajectory information.

6. The activity feature extraction device according to claim 5, wherein

the arithmetic device includes a feature extraction unit that extracts cluster attribute information of the cluster from facilities that the activity route data of the cluster passes through or stays at and transportation means used in the activity route data of the cluster.

7. The activity feature extraction device according to claim 6, wherein

the arithmetic device includes a person attribute input unit that acquires person attribute information of a user or a related person related to the user, and

the arithmetic device includes a belonging cluster calculation unit that calculates an attribute similarity

between the person attribute information and the cluster attribute information, and calculates belonging cluster information that indicates a belonging probability of the user or the related person to the cluster.

8. The activity feature extraction device according to claim 7, wherein

the arithmetic device includes a recommended spot extraction unit that extracts recommended spot information that matches an activity tendency of the user or the related person based on activity route data included in a cluster to which the user or the related person belongs.

9. The activity feature extraction device according to claim 8, wherein

the recommended spot extraction unit calculates a matching rate between the user or the related person and the recommended spot information based on activity route data included in the cluster to which the user or the related person belongs, and extracts the recommended spot information based on the calculated matching rate.

10. The activity feature extraction device according to claim 1, wherein

the arithmetic device includes an output device that

outputs an activity feature display screen that displays the activity route data and information of the cluster in a superimposed manner on map information and time information displayed on a screen.

11. The activity feature extraction device according to claim 6, wherein

an output device is provided that outputs an activity feature display screen that displays the activity route data and information of the cluster in a superimposed manner on map information and time information displayed on a screen, and further displays the cluster attribute information.

12. The activity feature extraction device according to claim 9, wherein

an output device is provided that outputs a spot recommendation screen that displays a location of the recommended spot information in a superimposed manner on map information displayed on a screen, and further displays a value determined based on the matching rate of the recommended spot information.

13. The activity feature extraction device according to claim 12, wherein

the recommended spot extraction unit calculates a

mutual matching rate based on both the matching rate of the user with the recommended spot information and the matching rate of the related person with the recommended spot information, and

the output device displays the calculated mutual matching rate on the spot recommendation screen.

14. An activity feature extraction method executed by an activity feature extraction device that analyzes activity route data in which person movement trajectory information and time information are associated with each other, wherein

the activity feature extraction device is configured with a computer including an arithmetic device that executes a predetermined arithmetic process and a storage device that the arithmetic device can access, and

the activity feature extraction method includes,

a spatial divergence calculation step in which the arithmetic device calculates a spatial divergence representing a spatial difference between the activity route data,

a temporal divergence calculation step in which the arithmetic device calculates a temporal divergence representing a temporal difference between the activity route data,

a spatiotemporal clustering step in which the arithmetic device calculates an activity feature similarity from the spatial divergence and the temporal divergence, and clusters the activity route data using the calculated activity feature similarity, and

an output generation step in which the arithmetic device outputs information on the cluster generated in the spatiotemporal clustering step.

15. An activity feature extraction device that analyzes activity route data in which person movement trajectory information and time information are associated with each other, wherein

the activity feature extraction device is configured with a computer including an arithmetic device that executes a predetermined arithmetic process and a storage device that the arithmetic device can access,

the arithmetic device includes a person attribute input unit that acquires person attribute information of a user or a related person related to the user,

the arithmetic device includes a belonging cluster calculation unit that calculates an attribute similarity between cluster attribute information of a cluster to which the activity route data is clustered and the person attribute information, and calculates belonging cluster

information that indicates a belonging probability of the user or the related person to the cluster, and

the arithmetic device includes a recommended spot extraction unit that extracts recommended spot information that matches an activity tendency of the user or the related person based on activity route data included in a cluster to which the user or the related person belongs.