

INFORMATION PROCESSING SYSTEM AND INFORMATION PROCESSING METHOD

CROSS-REFERENCE TO RELATED APPLICATION

[0001] The present international application claims priority based on Japanese Patent Application No. 2021-182537 filed to Japanese Patent Office on November 9, 2021 and Japanese Patent Application No. 2022-048893 filed to Japanese Patent Office on March 24, 2022, and the contents of Japanese Patent Application No. 2021-182537 and Japanese Patent Application No. 2022-048893 are incorporated by reference in their entirety.

TECHNICAL FIELD

[0002] The present disclosure relates to an information processing system and an information processing method.

BACKGROUND ART

[0003] Conventionally, customer purchasing behavior has been analyzed based on product sales data. Contact behavior of customers with mass media and network content has also been analyzed. Various types of information regarding customers are also collected by a questionnaire form or a face-to-face question form.

[0004] Data fusion technologies are also known that combine two or more data collected by different means based on common variables. In particular, a technology related to data fusion between a first data set including a first feature data for each customer for first customers and a second data set including a second feature data for each customer for second customers has already been disclosed (see, for example, Patent Document 1). The data fusion is performed to combine the first feature data and the second feature data of similar customers based on variables common between the first data

set and the second data set, for example, demographic attributes of the customer.

PRIOR ART DOCUMENTS

PATENT DOCUMENTS

[0005] Patent Document 1: Japanese Unexamined Patent Application
Publication No. 2016-126609

SUMMARY OF THE INVENTION

PROBLEMS TO BE SOLVED BY THE INVENTION

[0006] In the conventional data fusion technology, a common variable related to customers is required between a first data set and a second data set to be combined in order to determine similar customers by using the common variable. Therefore, data without the common variable cannot be combined.

[0007] Therefore, according to one aspect of the present disclosure, it is desirable to provide a technology capable of realizing association between first entities and second entities on the basis of a first data set related to the first entities and a second data set related to the second entities regardless of the common variable.

MEANS FOR SOLVING THE PROBLEMS

[0008] According to one aspect of the present disclosure, an information processing system is provided. The information processing system includes a first acquisition unit, a second acquisition unit, a dimension reduction unit, and an association unit. The first acquisition unit is configured to acquire a first data set related to first entities. The first data set may describe a feature of each of the first entities.

[0009] The second acquisition unit is configured to acquire a second data set related to second entities. The second data set may describe a feature of each of the second entities.

- [0010] The dimension reduction unit is configured to generate a group of first low-dimensional feature vectors corresponding to a group of first feature vectors and a group of second low-dimensional feature vectors corresponding to a group of second feature vectors by performing dimension reduction processing on the group of the first feature vectors specified from the first data set and the group of the second feature vectors specified from the second data set. The group of the second low-dimensional feature vectors may be a group of feature vectors having the same number of dimensions as the group of the first low-dimensional feature vectors.
- [0011] Each of the first feature vectors may represent a feature of a corresponding one of the first entities. Each of the second feature vectors may represent a feature of a corresponding one of the second entities.
- [0012] The association unit is configured to associate each of the first entities with at least one of the second entities on the basis of the group of the first low-dimensional feature vectors and the group of the second low-dimensional feature vectors.
- [0013] When a set of first entities and a set of second entities are subsets from a population common to each other or subsets from populations related to each other, even if there is no common variable between the first feature vector and the second feature vector, the feature of the first entity and the feature of the second entity can be expressed by a combination of components common to each other or related to each other by dimension reduction.
- [0014] That is, according to the dimension reduction, it is possible to extract key feature components common to or related to each other from the first feature vector and the second feature vector. Therefore, according to the comparison of the low-dimensional feature vectors, a conformance degree between the first entity and the second entity can be appropriately determined.
- [0015] Therefore, according to one aspect of the present disclosure, it is possible to appropriately perform association between the first entities and the second entities on the basis of the first data set related to the first entities

and the second data set related to the second entities regardless of the common variable.

[0016] According to one aspect of the present disclosure, the association unit can associate each of the first entities with at least one of the second entities on the basis of the similarity between the first entities specified from the group of the first low-dimensional feature vectors and the similarity between the second entities specified from the group of the second low-dimensional feature vectors such that the mutual relationship between the first entities regarding the similarity matches the mutual relationship between the second entities.

[0017] When the set of first entities and the set of second entities are subsets from a common population or subsets from mutually related populations, the mutual relationship regarding the similarity between the entities is approximately common or related to each other between the set of first entities and the set of second entities, similarly to the population(s).

[0018] Therefore, by associating each of the first entities with at least one of the second entities such that the mutual relationship between the first entities regarding the similarity matches the mutual relationship between the second entities, each of the first entities can be associated with an appropriate second entity having high identity or strong relationship.

[0019] According to one aspect of the present disclosure, a group of first low-dimensional feature vectors may be defined by a first feature space. A group of the second low-dimensional feature vectors may be defined by a second feature space.

[0020] The association unit may search a mapping for mapping the first entities on the first feature space to the second feature space such that a distribution of the the first entities in the first feature space specified from the group of the first low-dimensional feature vectors matches a distribution of the second entities in the second feature space specified from the group of the second low-dimensional feature vectors.

[0021] The association unit may be configured to associate each of the first entities with at least one of the second entities on the basis of mapping.

[0022] According to one aspect of the present disclosure, the association unit searches, as a matrix Ω^* , a matrix Ω that maximizes a value $Z(\Omega)$ according to the following formula including a matrix K, a matrix L, and a matrix H,

[0023] [Mathematical formula 1]

$$\begin{aligned}\Omega^* &= \operatorname{argmax}_{\Omega} Z(\Omega) \\ Z(\Omega) &= \operatorname{trace}(K' \Omega^T L' \Omega) \\ K' &= HKH \\ L' &= HLH\end{aligned}$$

[0024] and may be configured to associate each of the first entities with at least one of the second entities on the basis of the matrix Ω^* . T is a transpose symbol. trace is a diagonal sum of the matrix X.

[0025] The matrix K may be a matrix with N rows and N columns. The number of the first entities may be N. The number of the second entities may be the same as that of the first entity. The matrix K may be a first similarity matrix in which a value of an element in an i-th row and a j-th column represents similarity between an i-th entity and a j-th entity among the first entities.

[0026] The value of the element in the i-th row and the j-th column in the matrix K may be calculated based on the first low-dimensional feature vector of the i-th entity among the first entities and the first low-dimensional feature vector of the j-th entity among the first entities.

[0027] The matrix L may be a matrix with N rows and N columns. The matrix L is a second similarity matrix in which a value of an element in an i-th row and a j-th column represents similarity between an i-th entity and a j-th entity among the second entities.

[0028] The value of the element in the i-th row and the j-th column in the matrix L may be calculated based on the second low-dimensional feature vector of the i-th entity among the second entities and the second low-dimensional feature vector of the j-th entity among the second entities.

- [0029] The matrix H may be a matrix with N rows and N columns. The matrix H may be a matrix in which a value of an element in an i -th row and an j -th column indicates a value of $1-1/N$ when $i = j$ and indicates a value of 0 when $i \neq j$.
- [0030] According to one aspect of the present disclosure, the association unit may change the dimension reduction method in the dimension reduction processing based on the matrix Ω^* . For example, the association unit may change the dimension reduction method in the dimension reduction processing such that a distance in a feature space between the first low-dimensional feature vector and the second low-dimensional feature vector corresponding to each other in the group of the first low-dimensional feature vectors and the group of the second low-dimensional feature vectors is shortened.
- [0031] According to one aspect of the present disclosure, the association unit may be configured to improve the matrix Ω^* by repeatedly performing a re-search processing related to the matrix Ω^* until a predetermined condition is satisfied, and associate each of the first entities with at least one of the second entities on the basis of the improved matrix Ω^* .
- [0032] The re-search processing may include changing the dimension reduction method in the dimension reduction processing based on the matrix Ω^* . The re-search processing may include causing the dimension reduction unit to perform the dimension reduction processing based on dimension reduction method changed, and re-searching the matrix Ω^* on the basis of a group of the first low-dimensional feature vectors and a group of the second low-dimensional feature vectors newly obtained by the dimension reduction processing.
- [0033] According to the information processing system in which the association unit is configured as described above, association between the first entities and the second entities can be performed with high accuracy.
- [0034] According to one aspect of the present disclosure, the first data set may include two or more first feature data. Each of the two or more first feature

data may represent a feature of a corresponding one of the the first entities. The second data set may include two or more second feature data. Each of the two or more second feature data may represent a feature of a corresponding one of the second entities.

[0035] According to one aspect of the present disclosure, the information processing system may further include a data fusion unit. The data fusion unit may be configured to generate an extended data set by combining each of the two or more first feature data with one of the two or more second feature data, on the basis of the association by the association unit between the first entities and the second entities. The extended data set may include two or more extended data. Each of the two or more extended data may be a combined data of one of the two or more first feature data and a corresponding one of the two or more second feature data.

[0036] According to such an information processing system, it is possible to generate a data set having a large amount of information by combining two or more data sets.

[0037] According to one aspect of the present disclosure, the first entities each may be a person. The second entities each may be a person. The first data set may be a data set that describes a first feature of each person belonging to a first group. The second data set may be a data set describing a second feature of each person belonging to a second group.

[0038] It is considered that characteristics related to behavior, interest, and the like of a person greatly affect a demographic attribute, and a characteristic distribution according to the demographic attribute does not greatly change among different groups of people. Therefore, according to the information processing system according to one aspect of the present disclosure, it is possible to appropriately associate people between different groups without a common variable.

[0039] According to one aspect of the present disclosure, a combination of the first feature and the second feature may be a combination of a feature related

to purchasing behavior, and a feature related to movement in at least one of an online space and an offline space and/or a feature related to visits to two or more points in at least one of the online space and the offline space. The association of entities based on data sets related to such features, furthermore, data fusion is useful for human behavior analysis.

[0040] According to one aspect of the present disclosure, identification information of an information terminal corresponding to each of the second entities may be associated with the second data set.

[0041] According to one aspect of the present disclosure, the information processing system may include a selection unit that selects, as delivery destinations of information content, at least a part of a set of second entities associated by the association unit with any one of the first entities among the second entities.

[0042] According to one aspect of the present disclosure, the information processing system may include a delivery unit configured to deliver the information content to a set of information terminals corresponding to delivery destinations of the information content on the basis of the identification information.

[0043] This information processing system functions effectively when the first entities and the second entities are persons. According to the above delivery method, even when a relationship between the first entity and the information terminal is unknown, the information content can be appropriately delivered to the information terminal of the second entity corresponding to the first entity by utilizing the identification information of the information terminal associated with the second entity.

[0044] According to one aspect of the present disclosure, the selection unit may be configured to select, as delivery destinations of the information content, a first set that is a set of second entities associated by the association unit with any one of the first entities, and a second set having similar features to the first set among the second entities. According to such selection of the

delivery destinations, the delivery destinations can be extended within an appropriate range on the basis of the second data set, and the information content can be delivered.

[0045] According to one aspect of the present disclosure, the second data set may be a data set that describes features related to behavior of each of the second entities. In this case, the information processing system may include an estimation unit that calculates, for each of one or more entities of interest, an estimation value regarding the behavior of a corresponding entity of interest. The one or more entities of interest may be at least a part of the first entities. The estimation value may be calculated on the basis of a feature related to behavior of at least one of the second entities associated with the corresponding entity of interest. The first entities and the second entities may be persons.

[0046] According to the information processing system including the estimation unit described above, it is possible to estimate, through the second data set, behavior of the first entity that cannot be determined only by the first data set. The estimation may be a prediction.

[0047] According to one aspect of the present disclosure, an information processing method corresponding to the method performed by the information processing system described above may be provided. According to one aspect of the present disclosure, a computer-implemented information processing method may be provided. The information processing method may include acquiring a first data set related to the first entities, which describes a feature of each of the first entities.

[0048] The information processing method may include acquiring a second data set related to the second entities, which describes a feature of each of the second entities.

[0049] The information processing method may include generating a group of first low-dimensional feature vectors corresponding to a group of first feature vectors and a group of second low-dimensional feature vectors corresponding

to a group of second feature vectors by performing dimension reduction processing on the group of the first feature vectors specified from the first data set and the group of the second feature vectors specified from the second data set. The group of the second low-dimensional feature vectors may be a group of feature vectors having the same number of dimensions as the group of the first low-dimensional feature vectors.

[0050] Each of the first feature vectors may represent a feature of a corresponding one of the first entities. Each of the second feature vectors may represent a feature of a corresponding one of the second entities.

[0051] The information processing method may include associating each of the first entities with at least one of the second entities on the basis of a group of the first low-dimensional feature vectors and a group of the second low-dimensional feature vectors.

[0052] According to one aspect of the present disclosure, the associating may include associating each of the first entities with one of the second entities on the basis of the similarity between the first entities specified from the group of the first low-dimensional feature vectors and the similarity between the second entities specified from the group of the second low-dimensional feature vectors such that the mutual relationship between the first entities regarding the similarity matches the mutual relationship between the second entities.

[0053] According to such an information processing method, as in the above-described information processing system, association between the first entities and the second entities can be realized on the basis of the first data set related to the the first entities and the second data set related to the the second entities regardless of the common variable.

[0054] According to one aspect of the present disclosure, a computer program including instructions for causing a computer to perform the above-described information processing method may be provided. According to one aspect

of the present disclosure, a computer-readable, non-transitory, and tangible storage medium storing a computer program may be provided.

BRIEF DESCRIPTION OF THE DRAWINGS

[0055] FIG. 1 is a block diagram illustrating a configuration of an information processing system.

FIG. 2 is a flowchart illustrating analysis processing performed by a processor.

FIG. 3A is a diagram illustrating a configuration of a first data set, and FIG. 3B is a diagram illustrating a configuration of a second data set.

FIGS. 4A and 4B are diagrams for explaining a method of searching a matrix Ω .

FIG. 5 is a diagram illustrating a configuration of a correspondence table generated by a processor.

FIG. 6 is a diagram illustrating a configuration of an extended data set generated by a processor.

FIG. 7 is a flowchart illustrating analysis processing performed by a processor in a second embodiment.

FIG. 8 is a flowchart illustrating evaluation processing performed by a processor in a third embodiment.

FIG. 9 is a flowchart illustrating selection processing performed by the processor in the third embodiment.

FIG. 10 is a block diagram illustrating a configuration of a delivery system in a fourth embodiment.

FIG. 11 is a diagram illustrating a configuration of an internal data set in the fourth embodiment.

FIG. 12 is a flowchart illustrating delivery control processing performed by the processor in the fourth embodiment.

FIG. 13 is a flowchart illustrating delivery control processing performed by a processor in a fifth embodiment.

FIG. 14 is a flowchart illustrating prediction processing performed by a processor in a sixth embodiment.

EXPLANATION OF REFERENCE NUMERALS

[0056] 1...information processing system, 11, 31...processor, 13, 33...memory, 15, 35...storage, 15A...first data set, 15B...second data set, 15C...extended data set, 17...user interface, 19, 39...communication interface, 30...delivery system, 35A...external data set, 35B...internal data set, 35C...extended data set, 40...user company side system, Pr, Pr1...computer program

MODE FOR CARRYING OUT THE INVENTION

[0057] Exemplary embodiments of the present disclosure will be described below with reference to the drawings.

[0058] <First Embodiment>

An information processing system 1 of the present embodiment is configured by installing a dedicated computer program Pr in a general-purpose computer. As illustrated in FIG. 1, the information processing system 1 includes a processor 11, a memory 13, a storage 15, a user interface 17, and a communication interface 19.

[0059] The processor 11 performs processing according to a computer program Pr stored in the storage 15. The memory 13 is a primary storage device including a RAM, and is used as a work area when processing is performed by the processor 11.

[0060] The storage 15 is, for example, a secondary storage device including a hard disk drive or a solid-state drive, and stores the computer program Pr and various data used at the time of execution of processing according to the computer program Pr.

[0061] The user interface 17 includes an input device and a display. The input device is provided to input, to the processor 11, an operation signal from a user operating the information processing system 1. The display is provided

to display various types of information to the user. Examples of input devices include a keyboard and a pointing device.

[0062] The communication interface 19 includes a local area network (LAN) interface and a universal serial bus (USB) interface, and is used for communication with an external device. The information processing system 1 transmits and receives data to and from the external device through the communication interface 19.

[0063] The processor 11 in the information processing system 1 performs processing according to the computer program Pr to generate an extended data set 15C obtained by extending a first data set 15A acquired from the external device through the communication interface 19 using a second data set 15B.

[0064] The extended data set 15C is a data set obtained by adding information included in the second data set 15B to the first data set 15A. With the extension, the amount of information of each entity described by the first data set 15A increases. The entity is, for example, a person, in particular an individual. An increase in the information amount is performed for analyzing behavior of a person or delivering an advertisement based on the extended data set 15C.

[0065] Specifically, when an execution instruction from the user is input through the user interface 17, the processor 11 of the information processing system 1 performs analysis processing illustrated in FIG. 2. When the analysis processing illustrated in FIG. 2 is started, the processor 11 acquires the first data set 15A and the second data set 15B to be subjected to data fusion (S110, S120).

[0066] In S110 and S120, the processor 11 can read, from the storage 15, the first data set 15A and the second data set 15B stored in advance in the storage 15. As a result, the processor 11 can acquire the first data set 15A and the second data set 15B.

[0067] The first data set 15A and the second data set 15B to be acquired can be designated by the user. The user can collect the first data set 15A and the

second data set 15B to be subjected to data fusion in advance and store them in the storage 15.

[0068] Alternatively, the processor 11 can acquire the first data set 15A from a first external device and acquire the second data set 15B from a second external device, by communication using the communication interface 19.

[0069] The first data set 15A is a data set related to first entities, and is a data set describing a first feature of each of the first entities. The first data set 15A is a set of first feature data, and each of the first feature data represents a first feature of a corresponding one of the first entities.

[0070] The second data set 15B is a data set related to second entities, and is a data set describing a second feature of each of the second entities. The second feature may be a feature different from the first feature. Specifically, the second data set 15B is a set of second feature data, and each of the second feature data represents a second feature of a corresponding one of the second entities.

[0071] A set of the first entities and a set of the second entities are, for example, different subsets in a common population. The population may be a set of persons or a set of consumers. For example, the set of the first entities may be a set of persons corresponding to customers of a first company. For example, the set of the second entities may be a set of persons corresponding to customers of a second company different from the first company.

[0072] Alternatively, the set of the first entities may be a set of persons who are collection targets of a first behavior. The set of the second entities may be a set of persons who are collection targets of a second behavior.

[0073] A first data set 15A illustrated in FIG. 3A is data regarding a set of first persons, and includes a feature data regarding purchasing behavior for each person. Each of the feature data is associated with an ID of a corresponding person and indicates, for products P1, P2, P3, ..., whether each product has been purchased by the corresponding person by a binary of 1 or 0.

[0074] A second data set 15B illustrated in FIG. 3B is data regarding a set of second persons, and includes a feature data regarding browsing behavior of web content for each person. Each of the feature data is associated with an ID of a corresponding person and indicates, for web sites S1, S2, S3, ..., whether the corresponding person has visited each web site and browsed the web content by a binary of 1 or 0.

[0075] In S110, based on the first feature data for each of the first entities included in the acquired first data set 15A, the processor 11 generates an M1-dimensional feature vector $x = (x_1, x_2, x_3, \dots)$ for each of the first entities. According to an example, the elements $x_1, x_2, x_3, \dots$ of the feature vector x may represent the presence or absence of purchase of the products P1, P2, P3, ... of the corresponding person, respectively.

[0076] Similarly, in S120, based on the second feature data for each of the second entities included in the acquired second data set 15B, the processor 11 generates an M2-dimensional feature vector $y = (y_1, y_2, y_3, \dots)$ for each of the second entities. According to an example, the elements $y_1, y_2, y_3, \dots$ of the feature vector y may represent the presence or absence of browsing of web content on the websites S1, S2, S3, ... of the corresponding person, respectively.

[0077] Thereafter, the processor 11 performs dimension reduction processing (S130) on a group of the feature vectors x to convert each feature vector x from the M1-dimensional feature vector into a low-dimensional feature vector $Dx = (Dx_1, Dx_2, \dots)$ which is an M-dimensional feature vector smaller than the M1-dimensional feature vector. As a result, the processor 11 generates a group of low-dimensional feature vectors Dx corresponding to the group of feature vectors x . A lower right region of FIG. 3A illustrates an example of the low-dimensional feature vector Dx by a table.

[0078] The processor 11 further performs dimension reduction processing (S140) on a group of the feature vectors y to convert each feature vector y from the M2-dimensional feature vector into a low-dimensional feature vector

$D_y = (D_{y1}, D_{y2}, \dots)$ which is an M-dimensional feature vector smaller than the M2-dimensional feature vector. As a result, the processor 11 generates a group of low-dimensional feature vectors D_y corresponding to the group of feature vectors y . The low-dimensional feature vector D_y is a feature vector having the same number of dimensions M as the low-dimensional feature vector D_x . A lower right region of FIG. 3B illustrates an example of the low-dimensional feature vector D_y by a table.

[0079] As examples of an algorithm for realizing mapping to a low dimensional space, nonnegative matrix factorization, latent dirichlet allocation, singular value decomposition, and probabilistic latent semantic analysis are known. The dimension reduction processing in S130 and S140 can be performed using one of these algorithms.

[0080] According to the algorithm described above, the feature vectors may be dimensioned down so that key feature components that strongly characterize individual entities are extracted. Alternatively, the feature vector may be dimensioned down in a form with less loss of information for distinguishing each entity.

[0081] Thereafter, the processor 11 performs alignment processing of calculating a relationship between the first entities and the second entities on the basis of the group of the low-dimensional feature vectors D_x and the group of the low-dimensional feature vectors D_y (S150 to S180).

[0082] The alignment processing is performed using a technique of Kernelized Sorting. Hereinafter, details of the alignment processing using Kernelized Sorting will be described. However, the alignment processing may be realized by using adversarial learning, a gromov-wasserstein alignment technology, or an unbalanced optimal transport technology.

[0083] In S150, the processor 11 generates a similarity matrix K related to the first entities using the group of the low-dimensional feature vectors D_x . The similarity matrix K is a square matrix with N rows and N columns. Here, N

is the number of low-dimensional feature vectors D_x , in other words, the number of first entities.

[0084] The similarity matrix K is defined as a matrix in which the value K_{ij} of the element in an i -th row and an j -th column represents similarity between an i -th entity and a j -th entity in the first entities.

[0085] That is, the similarity matrix K is defined as a matrix that describes distribution of similarity between entities with respect to the first entities. In other words, the similarity matrix K is defined as a matrix describing the distribution of the entities on the feature space using a measure of the closeness between the entities with respect to the first entities.

[0086] Specifically, the similarity is calculated as a value $k(D_x[i], D_x[j])$ obtained by substituting a low-dimensional feature vector $D_x[i]$ that is the low-dimensional feature vector D_x of the i -th entity and a low-dimensional feature vector $D_x[j]$ that is the low-dimensional feature vector D_x of the j -th entity into a kernel function $k(a, b)$. That is, $K_{ij} = k(D_x[i], D_x[j])$.

[0087] Examples of the kernel function $k(a, b)$ include a Gaussian RBF (radial basis function) kernel expressed by the following formula. The similarity calculated using the kernel function $k(a, b)$ takes a value in a range from a value 0 to a value 1.

[0088] [Mathematical formula 2]

$$k(a, b) = \exp\left(\frac{-\|a - b\|^2}{\sigma^2}\right)$$

[0089] According to the kernel function $k(a, b)$, the value K_{ij} of the element of the similarity matrix K satisfies $0 < K_{ij} \leq 1$.

[0090] In S160, the processor 11 generates a similarity matrix L related to the second entities using the group of the low-dimensional feature vectors D_y . The similarity matrix L is a square matrix with N rows and N columns. Here, N is the number of low-dimensional feature vectors D_y , in other words, the

number of second entities. That is, the number of first entities and the number of second entities are the same.

[0091] Similarly to the similarity matrix K, the similarity matrix L is defined as a matrix in which the value L_{ij} of the element in the i-th row and the j-th column represents the similarity between the i-th entity and the j-th entity in the second entities. That is, the value L_{ij} of the element in the i-th row and the j-th column is k ($Dy [i]$, $Dy [j]$).

[0092] In subsequent S170, the processor 11 searches a matrix Ω that maximizes a value $Z(\Omega)$ according to the following formula, as a matrix Ω^* , using the similarity matrix K and the similarity matrix L.

[0093] [Mathematical formula 3]

$$\begin{aligned}\Omega^* &= \operatorname{argmax}_{\Omega} Z(\Omega) \\ Z(\Omega) &= \operatorname{trace}(K' \Omega^T L' \Omega) \\ K' &= HKH \\ L' &= HLH\end{aligned}$$

[0094] Here, the matrix H is a matrix with N rows and N columns, and is a diagonal matrix in which the value of an element in the i-th row and the j-th column indicates a value of $1-1/N$ when $i = j$ and indicates a value of 0 when $i \neq j$. T is a transpose symbol. $\operatorname{trace}(X)$ is a diagonal sum of the matrix X. The similarity matrices K and L are symmetric matrices. When an ideal Ω is found such that the matrix $\Omega^T L' \Omega$ is a transposed matrix of the matrix K', the value $Z(\Omega)$ is maximized.

[0095] Searching the matrix Ω^* corresponds to associating each of the first entities with at least one of the second entities on the basis of the similarity between the first entities specified from the group of the low-dimensional feature vectors Dx and the similarity between the second entities specified from the group of the low-dimensional feature vectors Dy such that the mutual relationship between the first entities regarding the similarity matches the mutual relationship between the second entities.

[0096] In other words, searching the matrix Ω^* corresponds to searching a mapping for mapping the first entities on the first M-dimensional feature space to the second M-dimensional feature space, so that a distribution of the first entities in the first M-dimensional feature space specified from the group of the low-dimensional feature vectors D_x , the distribution of the first entities defined by the similarity between the entities, matches a distribution of the second entities in the second M-dimensional feature space specified from the group of the low-dimensional feature vectors D_y .

[0097] The left graph of FIG. 4A conceptually represents the distribution of the first entities, and the left graph of FIG. 4B conceptually represents the distribution of the second entities. In the examples illustrated in FIGS. 4A and 4B, two-dimensional low-dimensional feature vectors D_x and D_y are defined only for technical description. Each of points denoted by reference numerals E11, E12, E13, E14, E15, E16, and E17 indicates a position of each of the first entities on the feature space. Each of points denoted by reference numerals E21, E22, E23, E24, E25, E26, and E27 indicates a position of each of the second entities on the feature space.

[0098] As can be understood from FIG. 4B, according to this example, the component D_{y1} of the low-dimensional feature vector D_y corresponds to the component D_{x2} of the low-dimensional feature vector D_x , and the component D_{y2} of the low-dimensional feature vector D_y corresponds to the component D_{x1} of the low-dimensional feature vector D_x .

[0099] That is, according to the example illustrated in Fig. 4A, the group of the first entities and the group of the second entities are only defined such that the arrangement and the order of dimensions for the entities are different between the similarity matrix K and the similarity matrix L , and substantially indicate the similarity distribution of the same group of entities.

[0100] When the group of the first entities and the group of the second entities have a collective property that is common or related to each other due to the same population or the like, it is possible to extract a feature component

essentially common to each entity even if there is no common variable between the first data set 15A and the second data set 15B of the information source by reducing the dimensionality of the feature vectors x and y .

[0101] However, even with such a reduction in dimension, the low-dimensional feature vectors Dx and Dy only have the same feature components, and the arrangement of the feature components cannot be aligned. In addition, the arrangement of entities is not uniform between the first data set 15A and the second data set 15B.

[0102] The search of the matrix Ω^* corresponds to an operation of searching the correspondence relationship between the irregular feature vectors Dx and Dy with respect to the arrangement of the entities and the arrangement of the dimensions, using the identity of the similarity distribution as a clue.

[0103] In subsequent S180, the processor 11 associates each of the first entities with at least one of the second entities, based on the matrix Ω^* . The element value in the i -th row and the j -th column in the matrix Ω^* represents a degree or a possibility that the i -th entity in the first entities and the j -th entity in the second entities correspond to each other, according to the distribution of the similarity.

[0104] Each element of the matrix Ω^* ideally takes 0 or 1, the sum of the element values in one row is 1 for each row, and the sum of the element values in one column is 1 for each column. When the matrix Ω^* is such an ideal matrix, the first entity of the row number and the second entity of the column number of the element whose value is 1 correspond to each other.

[0105] That is, when the element of the i -th row and the j -th column in the matrix Ω^* has a value of 1, it indicates that the i -th entity in the first entities and the j -th entity in the second entities correspond to each other.

[0106] However, in numerical calculation, the matrix Ω^* rarely becomes such an ideal matrix. Therefore, in S180, each of the first entities is associated with at least one of the second entities, by any one of the following methods.

- [0107] (Method 1) In the i-th row of the matrix Ω^* , an element having the largest value is searched. In a case where the element having the largest value is in the c-th column, the i-th entity in the first entities is associated with the c-th entity in the second entities. This is done for all rows.
- [0108] In this method, there is a possibility that two or more first entities is associated with one of the second entities. In order to suppress this possibility, a neighborhood search may be performed. As an example of the neighborhood search, Contextual Dissimilarity measure is known.
- [0109] (Method 2) In order to perform strict one-to-one association, each of the first entities is associated with non-overlapping one of the second entities by solving an optimal allocation problem using the matrix Ω^* as an input.
- [0110] In S180, the processor 11 may further output the correspondence table illustrated in FIG. 5 as a table for describing the correspondence between the first entities and the second entities. That is, the correspondence table may describe, in association with an ID of each of the first entities, an ID of a corresponding second entity. The correspondence table may be output and stored in the storage 15.
- [0111] Furthermore, the processor 11 performs data fusion processing (S190). In the data fusion processing, the processor 11 combines the first data set 15A and the second data set 15B, as a result of the association or on the basis of the correspondence table, to generate the extended data set 15C.
- [0112] The extended data set 15C includes two or more extended data. As illustrated in FIG. 6, each of the extended data is a combined data of one of the first feature data and a corresponding one of the second feature data.
- [0113] That is, the processor 11 generates the extended data set 15C by combining each of the first feature data included in the first data set 15A with one of the second feature data included in the second data set 15B based on the correspondence table.
- [0114] When the i-th entity in the first entities and the j-th entity in the second entities are associated with each other by the correspondence table, the

processor 11 combines a first feature data describing the feature of the i-th entity in the first entities and a second feature data describing the feature of the j-th entity in the second entities to generate the extended data of the i-th entity.

[0115] The extended data set 15C generated in this manner is stored in the storage 15. The extended data set 15C stored in the storage 15 is transferred to another system through the communication interface 19 on the basis of, for example, an instruction from the user input through the user interface 17.

[0116] Another system may be, for example, an advertisement delivery system. The advertisement delivery system can determine an entity of an advertisement delivery destination based on the extended data set 15C and deliver the advertisement to the entity.

[0117] When the data fusion processing is ended in S190, the processor 11 ends the analysis processing illustrated in Fig. 2.

[0118] As described above, according to the information processing system 1 of the present embodiment, even if there is no common variable between the first data set 15A and the second data set 15B, it is possible to appropriately associate the first entities and the second entities on the basis of the distribution of the similarity.

[0119] For proper association, it is preferable that a distribution of similarity between the first entities and the second entities be matched, similar, or related to each other.

[0120] Such a preferred condition is approximately satisfied when the set of first entities and the set of second entities are subsets from the same population. Therefore, when the first entities and the second entities are persons, that is, when data sets representing features related to persons are handled as the first data set 15A and the second data set 15B, the technology of the present embodiment effectively functions.

[0121] In particular, the behavior of a person often shows a tendency depending on a demographic attribute in particular. Therefore, when the first data set

15A and the second data set 15B are data sets based on collected data from groups estimated to have distributions of demographic attributes similar to each other, appropriate association between entities can be realized.

[0122] For example, even when the first data set 15A and the second data set 15B are data sets describing features of persons belonging to different populations having no common variable or data sets describing features of different behaviors, association between entities can be appropriately performed. Therefore, a data set useful for human psychological/behavior analysis can be generated as the extended data set 15C.

[0123] According to the example described above, the first data set 15A is a data set describing features related to purchasing behavior of each person belonging to the first group, and the second data set 15B is a data set describing features related to website visiting behavior and/or web content browsing behavior of each person belonging to the second group.

[0124] According to another example, for one of the first data set 15A and the second data set 15B, a data set describing features related to media contact behavior of a person, such as TV viewing behavior, may be used. For one of the first data set 15A and the second data set 15B, a data set describing a feature related to a usage status of a mobile terminal such as a smartphone may be used.

[0125] For one of the first data set 15A and the second data set 15B, a data set describing a feature related to movement of a person in an offline space (that is, a real space) may be used. The data set may describe, as characteristics related to movement of a person in the offline space, for example, characteristics related to visits to two or more places, movement paths, and/or movement means.

[0126] For one of the first data set 15A and the second data set 15B, a data set describing a feature related to movement of a person in an online space may be used. The data set may describe, as characteristics related to movement of a person in an online space, characteristics related to movement of a person

or net-surfing in a virtual reality (VR) space. For one of the first data set 15A and the second data set 15B, a data set based on data collected by a questionnaire may be used.

[0127] As the combination of the first data set 15A and the second data set 15B, a combination of a data set collected by a questionnaire and a data set related to TV viewing behavior, or a combination of a data set related to movement history and a data set related to purchase may be adopted.

[0128] In the above embodiment, processing such as ZCA whitening, normalization, and standardization may be performed on a group of the low-dimensional feature vectors D_x and D_y .

[0129] In the above embodiment, the number of dimensions M of the low-dimensional feature vectors D_x and D_y is determined by a designer or a user, but the information processing system 1 may be configured to search an optimum number of dimensions M . For example, the information processing system 1 may be configured to repeatedly perform the analysis processing illustrated in FIG. 2 while changing the number of dimensions M for the same data sets 15A and 15B, and automatically select the optimum number of dimensions M using a maximum value of $Z(\Omega)$ as an index.

[0130] <Second Embodiment>

An information processing system 1 of the second embodiment is configured such that a processor 11 performs analysis processing illustrated in FIG. 7 instead of the analysis processing illustrated in FIG. 2. Hereinafter, details of the analysis processing performed by the processor 11 will be selectively described as a description of the second embodiment. The configuration of the information processing system 1 not mentioned in the present embodiment may be understood to be the same as that of the first embodiment.

[0131] When the analysis processing illustrated in FIG. 7 is started, the processor 11 acquires the first data set 15A and the second data set 15B to be subjected to data fusion (S310, S320), similarly to the first embodiment.

[0132] Similarly to the processing in S110, the processor 11 generates the feature vector x for each first entity on the basis of the first data set 15A (S310). Similarly to the processing in S120, the processor 11 generates the feature vector y for each second entity on the basis of the second data set 15B (S320).

[0133] Furthermore, similarly to the processing in S130 and S140, the processor 11 generates a group of low-dimensional feature vectors Dx corresponding to the group of feature vectors x by the dimension reduction processing, and generates a group of low-dimensional feature vectors Dy corresponding to the group of feature vectors y (S330).

[0134] In subsequent S340, the processor 11 performs processing similar to the processing in S150, S160, and S170. That is, the processor 11 generates a similarity matrix K related to the first entities using the group of the low-dimensional feature vectors Dx , and generates a similarity matrix L related to the second entities using the group of the low-dimensional feature vectors Dy .

[0135] Furthermore, the processor 11 searches, as a matrix Ω^* , a matrix Ω that maximizes the value $Z(\Omega)$ described in the first embodiment, using the similarity matrix K and the similarity matrix L (S340). Here, the searched matrix Ω^* is expressed as a correspondence matrix Ω^* .

[0136] Thereafter, the processor 11 determines whether or not the repetition end condition is satisfied (S350). When determining that the repetition end condition is not satisfied (No in S350), the processor 11 performs processing of S360.

[0137] In S360, the processor 11 searches for a dimension reduction method that minimizes the cost of the Gromov-Wasserstein distance while fixing the correspondence matrix Ω^* searched in S340.

[0138] The state in which the correspondence matrix Ω^* is fixed corresponds to a state in which the correspondence between the first entities and the second entities is fixed. Searching the matrix Ω that maximizes, as the correspondence matrix Ω^* , the value $Z(\Omega)$ as described above corresponds to

searching a mapping for mapping the first entities on the first feature space to the second feature space so as to match the distribution of the second entities in the second feature space.

[0139] The cost of the Gromov-Wasserstein distance corresponds to the transport cost in the optimal transport problem between the first entities and the second entities when mapping the first entities into the second feature space.

[0140] The cost of the Gromov-Wasserstein distance can be calculated using the similarity matrices K and L and the correspondence matrix Ω^* . As described above, the similarity matrix K is a matrix including, as elements, the similarity between the first entities calculated on the basis of the dimensionally reduced low-dimensional feature vector Dx . The similarity matrix L is a matrix including, as elements, the similarity between the second entities calculated on the basis of the dimensionally reduced low-dimensional feature vector Dy .

[0141] Searching a dimension reduction method that minimizes the cost of the Gromov-Wasserstein distance corresponds to searching a dimension reduction method to generate low-dimension feature vectors Dx , Dy that best justify the correspondence between the first entities and the second entities indicated by the correspondence matrix Ω^* .

[0142] The minimization of the cost corresponds to searching a dimension reduction method according to the correspondence matrix Ω^* so that a distance on the feature space between the first entity and the second entity corresponding to each other, in other words, a distance on the feature space between the low-dimensional feature vector Dx of the first entity and the low-dimensional feature vector Dy of the second entity is shortened.

[0143] For example, in a case where the $M1$ -dimensional feature vector x is converted into the M -dimensional low-dimensional feature vector Dx , a conversion matrix Tx with M rows and $M1$ columns is applied to the feature vector x . In a case where the $M2$ -dimensional feature vector y is converted

into the M-dimensional low-dimensional feature vector D_y , a conversion matrix T_y with M rows and M_2 columns is applied to the feature vector y . At this time, the number of parameters m constituting the conversion matrices T_x and T_y is $(M * M_1 + M * M_2)$.

[0144] The search of the dimension reduction method is realized by, for example, searching a parameter m that minimizes the above-described cost as the parameter m of the conversion matrices T_x and T_y using a gradient method or the like.

[0145] Thereafter, the processor 11 reduces the dimensions of the feature vectors x and y by the searched dimension reduction method (for example, conversion matrices T_x and T_y), and calculates new low-dimensional feature vectors D_x and D_y (S370).

[0146] Using the similarity matrix K based on the new low-dimensional feature vector D_x and the similarity matrix L based on the new low-dimension feature vector D_y , the processor 11 searches for a matrix Ω that maximizes the value $Z(\Omega)$ as a new correspondence matrix Ω^* (S340).

[0147] By repeatedly performing the processing of S360, S370, and S340 in this manner, the processor 11 re-searches for the correspondence matrix Ω^* with high matching accuracy together with a better dimension reduction method.

[0148] When the repetition end condition is satisfied (Yes in S350), the processor 11 performs the processing of S380. The repetition end condition is satisfied when, for example, the processing of S340 is performed a predetermined number of times, or the amount of change in the correspondence matrix Ω^* by the re-search is less than a certain value.

[0149] In S380, the processor 11 associates each of the first entities with at least one of the second entities on the basis of the correspondence matrix Ω^* calculated at the end of the repetition processing, similarly to the processing of S180 in the first embodiment. The processor 11 may further store and output a correspondence table describing a correspondence between the first entities and the second entities.

[0150] Thereafter, the processor 11 performs the data fusion processing similarly to the processing in S190 to combine the first data set 15A and the second data set 15B to generate the extended data set 15C, and stores the generated extended data set 15C in the storage 15 (S390).

[0151] The information processing system 1 of the second embodiment described above can perform association between the first entities and the second entities with higher accuracy by the above-described repetition processing. Therefore, it is possible to generate the extended data set 15C with high accuracy.

[0152] <Third Embodiment>

 The information processing system 1 of the third embodiment is configured such that the processor 11 performs evaluation processing illustrated in FIG. 8 on the basis of an execution instruction from the user through the user interface 17. Hereinafter, details of the evaluation processing performed by the processor 11 will be described as a description of the third embodiment. The configuration of the information processing system 1 not mentioned in the present embodiment may be understood to be the same as that of the first or the second embodiment.

[0153] The evaluation processing is performed to evaluate whether or not the data set to be evaluated is an excellent data set for which association and data fusion in the analysis processing illustrated in FIG. 2 or FIG. 7 can be performed with high accuracy. The data set to be evaluated corresponds to a data set that may be used as the first data set 15A or the second data set 15B in the analysis processing.

[0154] When starting the evaluation processing, the processor 11 acquires a data set to be evaluated, which is designated together with an execution instruction from the user (S410). The processor 11 can acquire, from the storage 15, the designated data set to be evaluated.

[0155] Thereafter, the processor 11 generates a first feature vector x_1 and a second feature vector x_2 for each entity on the basis of the data set to be

evaluated (S420). The data set to be evaluated may include, for each entity, a feature data representing a feature of a corresponding entity by $(Q1 + Q2)$ pieces of elements.

[0156] The processor 11 can divide the $(Q1 + Q2)$ pieces of elements into a first element group including $Q1$ pieces of elements and a second element group including $Q2$ pieces of elements. Each of the $(Q1 + Q2)$ pieces of elements may be randomly classified into one of the first element group and the second element group.

[0157] Based on the data set to be evaluated, the processor 11 can generate, for each entity, a first feature vector x_1 describing a feature related to the first element group of a corresponding entity and a second feature vector x_2 describing a feature related to the second element group of the corresponding entity.

[0158] For example, in a case where the data set to be evaluated includes, for each entity, a feature data of the number of elements $Q = (Q1 + Q2)$ in which the feature vector $v = (v [1], v [2], v [3], \dots, v [Q])$ may be generated in S110, S120, S310, or S320, a first feature vector $x_1 = (v [1], v [2], \dots, v [Q1])$ including $Q1$ pieces of elements and a second feature vector $x_2 = (v [Q1 + 1], v [Q1 + 2], \dots, v [Q1 + Q2])$ including $Q2$ pieces of elements may be generated.

[0159] The first feature vector x_1 corresponds to a feature vector x for each entity in first entities, and the second feature vector x_2 corresponds to a feature vector y for each entity in the same second entities as the first entities.

[0160] Thereafter, the processor 11 performs processing similar to the processing performed in S130 to S170 on the first feature vector x_1 and the second feature vector x_2 in S430 and S440.

[0161] In S430, similarly to the processing in S130 and S140, the processor 11 performs the dimension reduction processing on the first feature vector x_1 for each of the first entities and the second feature vector x_2 for each of the

second entities to generate a low-dimensional feature vector Dx_1 and a low-dimensional feature vector Dx_2 having the same number of dimensions.

[0162] Based on the low-dimensional feature vector Dx_1 of each of the first entities, the processor 11 generates a similarity matrix representing the similarity of the low-dimensional feature vectors Dx_1 between the first entities corresponding to the similarity matrix K . Based on the low-dimensional feature vector Dx_2 of each of the second entities, the processor 11 further generates a similarity matrix representing the similarity of the low-dimensional feature vectors Dx_2 between the second entities corresponding to the similarity matrix L .

[0163] Based on these similarity matrices, the processor 11 searches for a matrix Ω that maximizes the value $Z(\Omega)$ as a correspondence matrix Ω^* (S440).

[0164] Thereafter, the processor 11 calculates, as a score, a degree to which the correspondence matrix Ω^* correctly represents a correspondence between the first entities and the second entities for the first entities corresponding to the group of the low-dimensional feature vectors Dx_1 and the second entities corresponding to the group of the low-dimensional feature vectors Dx_2 (S450).

[0165] As a result, the processor 11 evaluates whether or not the data set to be evaluated is an excellent data set for which association and data fusion by the analysis processing can be performed with high accuracy (S450).

[0166] The processor 11 may store a correct correspondence between the first entities and the second entities when generating the feature vector x_1 for each first entity and the feature vector x_2 for each second entity in advance in S420.

[0167] The processor 11 performs processing similar to the analysis processing in S430 and S440 in an environment in which a correct answer of the correspondence is stored in this manner to calculate the correspondence

matrix Ω^* , and compares the correspondence specified from the correspondence matrix Ω^* with the correct answer.

[0168] For example, the processor 11 performs processing of associating each of the first entities with one of the second entities on the basis of the correspondence matrix Ω^* , similarly to the processing in S180 and S380.

[0169] The processor 11 determines that the association has succeeded when the first entity and the second entity associated on the basis of the correspondence matrix Ω^* are the same entity in the data set to be evaluated, and determines that the association has failed when the first entity and the second entity are not the same entity.

[0170] The processor 11 may calculate a rate of successful association among the entire entities as a score of the data set to be evaluated (S450). Thereafter, the processor 11 outputs the calculated score as an evaluation result (S460), and ends the evaluation processing.

[0171] In a case where association and data fusion based on one data set cannot be performed with high accuracy, it can be estimated that the data set does not have sufficient information or data structure to realize highly accurate association and data fusion with respect to the features of the set.

[0172] This lack of information also affects the accuracy in a case where analysis processing is performed on two different data sets to perform association and data fusion. Therefore, according to the evaluation processing, it can be estimated in advance whether the data set to be evaluated is a data set for which data fusion without a common variable can be performed with high accuracy.

[0173] In S460, the processor 11 can transmit whether or not the data set to be evaluated is an excellent data set to the user of the information processing system 1 by outputting the score. As a result, the user can obtain the highly reliable extended data set 15C by adopting an appropriate combination of the first data set 15A and the second data set 15B for the analysis processing.

[0174] In order to obtain the desired extended data set 15C, an environment is conceivable in which it is sufficient to adopt any of data sets similar to each other as the first data set 15A coupled to the second data set 15B.

[0175] For example, it is considered that a first data set 15A regarding purchasing behavior and a second data set 15B regarding website visiting behavior/web content browsing behavior are combined to generate an extended data set 15C. In this case, it is considered to be sufficient to generate the extended data set 15C using the data set related to the purchasing behavior of the customer of any one of distribution organizations as the first data set 15A.

[0176] Examples of the distribution organizations include convenience store chains. The data set related to the purchase of each convenience store chain may include information on the same type of purchasing behavior as that of another convenience store chain as the purchase behavior of consumers.

[0177] Therefore, it is considered to be sufficient to generate the extended data set 15C using the data set related to the purchasing behavior of the customer for any one of the convenience store chains, as the first data set 15A.

[0178] In a case where two or more data sets exist as candidates of the first data set 15A (or the second data set 15B), the evaluation processing described above can be used to select, from the two or more data sets, an optimal data set as the first data set 15A (or the second data set 15B), from the viewpoint of the accuracy of association and data fusion.

[0179] For example, the processor 11 can adopt one candidate, as the data set to be subjected to data fusion, from two or more candidates of the data set to be subjected to data fusion by performing selection processing illustrated in FIG. 9 as necessary in any processing of S110, S120, S310, and S320. The data set to be subjected to data fusion in S110 and S310 corresponds to the first data set 15A, and the data set to be subjected to data fusion in S120 and S320 corresponds to the second data set 15B.

- [0180] When the selection processing illustrated in FIG. 9 is started, the processor 11 acquires two or more data sets as two or more candidates of the data set to be subjected to data fusion (S510). The processor 11 can acquire two or more data sets designated by the user from the storage 15.
- [0181] Thereafter, the processor 11 sets one of the data sets as the data set to be evaluated (S520), and performs the evaluation processing illustrated in FIG. 8 (S530). Until the processor 11 performs the evaluation processing on all of the data sets (Yes in S540), for each of the data sets, a corresponding data set is set as the data set to be evaluated (S520), and the performing the evaluation processing (S530) is repeated. As a result, the score calculated in S450 is acquired for each of the data sets.
- [0182] When the evaluation processing is performed on all of the data sets and the scores are acquired (Yes in S540), the processor 11 adopts, as the data set to be subjected to data fusion, the data set having the highest score among the data sets (S550). Thereafter, the selection processing ends. In S110, S120, S310, and S320, the processor 11 can generate feature vectors (x or y) based on the adopted data set to be subjected to data fusion.
- [0183] By performing the selection processing in this manner and selecting an optimal data set from among the two or more candidates, it is possible to generate the extended data set 15C with high accuracy.
- [0184] In addition, in the example of the purchasing behavior, the two or more candidates of the data set to be subjected to data fusion may include two or more data sets representing the purchasing behavior of consumers with different parameters. For example, the first candidate may be a data set for which a feature vector including, as an element, the number of pieces purchased for each product can be generated for each consumer as an entity. The second candidate may be a data set for which a feature vector including, as an element, a purchase amount for each product can be generated for each consumer as an entity.

[0185] Preparing two or more data sets describing the same type of features with different parameters and selecting a data set suitable for data fusion leads to generation of a better extended data set 15C.

[0186] <Fourth Embodiment>

A delivery system 30 of a fourth embodiment illustrated in FIG. 10 is a system that combines an external data set 35A that is a data set provided from the outside of the delivery system 30 and an internal data set 35B that is a data set held inside the delivery system 30 using the data fusion technology of the first embodiment or the second embodiment, and performs advertisement delivery on the basis of an extended data set 35C generated by the combination.

[0187] As illustrated in FIG. 10, the delivery system 30 includes a processor 31, a memory 33, a storage 35, and a communication interface 39. The processor 31 performs processing according to a computer program Pr1 stored in the storage 35. The storage 35 further includes the internal data set 35B.

[0188] As illustrated in FIG. 11, the internal data set 35B includes, for each user, a feature data describing a feature of online behavior of a corresponding user in association with an advertisement ID of the corresponding user. As is well known, the advertisement ID is an identification code used for advertisement, and is an ID unique to an information terminal.

[0189] The feature data associated with the advertisement ID describes the feature of the online behavior of the user observed through the information terminal to which the corresponding advertisement ID is assigned. The online behavior includes a website visiting behavior and a web content browsing behavior.

[0190] The delivery system 30 is connected to a wide area network through the communication interface 39, and provides an advertisement delivery service via the wide area network. A user company side system 40, which is a system of a company that uses the advertisement delivery service, provides the delivery system 30 with delivery designation information together with

advertisement content to be delivered. The advertisement content is information content for advertisement. The delivery designation information includes target designation information for designating a delivery target and information on designation of number of deliveries for designating the number of deliveries.

[0191] The user company side system 40 further provides the delivery system 30 with, as the external data set 35A, a customer data set that is a data set describing features of customers corresponding to delivery destination candidates.

[0192] The customer data set may be, for example, a data set describing features related to purchasing behavior of customers who use stores operated by a user company. For example, the customer data set may include, as a feature data for each customer, a feature data describing a purchase amount for each product of a corresponding customer regarding two or more products.

[0193] When a delivery request is input from the user company side system 40 through the communication interface 39, the processor 31 performs delivery control processing illustrated in FIG. 12 on the basis of the computer program Pr1.

[0194] When the delivery control processing is started, the processor 31 acquires, from the user company side system 40, the delivery designation information including the target designation information and the information on designation of number of deliveries together with the advertisement content to be delivered, and the customer data set as the external data set 35A (S610).

[0195] Thereafter, the processor 31 uses the external data set 35A as the first data set 15A and further uses the internal data set 35B as the second data set 15A, and performs processing similar to the processing of S110 to S190 in the analysis processing. As a result, the processor 31 combines the external data set 35A and the internal data set 35B to generate the extended data set 35C (S620).

[0196] By combining the external data set 35A and the internal data set 35B, the feature data for each customer included in the external data set 35A is associated with the advertisement ID of the user who is likely to be the same person as the customer included in the internal data set 35B.

[0197] The extended data set 35C includes, for each entity, an extended data in which a feature data included in the external data set 35A of a corresponding customer and a feature data included in the internal data set 35B of a corresponding user are combined. Each extended data is associated with an advertisement ID of a corresponding user included in the internal data set 35B.

[0198] The entity referred to herein is a combination of a customer and a user associated with each other by data fusion. In the data fusion, the customer and the user are associated on a one-to-one basis. For example, the extended data set 35C may be a data set having a structure in which an advertisement ID of each entity is described in the column having “ID2_1”, “ID2_2”, and “ID2_3” illustrated in the extended data set 15C illustrated in FIG. 6.

[0199] Thereafter, the processor 31 calculates a score regarding the possibility that each entity in the extended data set 35C is a delivery target (S630). For example, in a case where the external data set 35A is a data set related to the purchasing behavior of the customers and the internal data set 35B is a data set related to the online behavior of the users, the processor 31 inputs a feature data related to the purchasing behavior and a feature data related to the online behavior of each entity in the extended data set 35C to a predetermined function, and calculates a score obtained by quantifying the possibility that a corresponding entity is a delivery target.

[0200] The delivery targets are a group of delivery destination consumers narrowed down by parameters characterizing consumers such as gender, age, purchasing tendency, online behavior tendency, interest, and concern, and is designated through target designation information.

[0201] After the score is calculated in S630, the processor 31 determines entities corresponding to the number of deliveries designated from the user

company side system 40 as the delivery destinations of content in descending order of the calculated score among a group of entities associated with the advertisement ID (S640). In this way, the processor 31 selects, as delivery destinations of the advertisement content, at least some of two or more users corresponding to the internal data set 35B associated with any of two or more customers corresponding to the external data set 35A.

[0202] Thereafter, the processor 31 transmits the advertisement content provided from the user company side system 40 to information terminals of the determined delivery destinations through the wide area network (S650). The advertisement content is delivered to the information terminals identified from the advertisement IDs of the delivery destinations. Thereafter, the processor 31 ends the delivery control processing.

[0203] According to the delivery system 30 of the fourth embodiment described above, the advertisement ID can be associated with the feature data of the customer whose advertisement ID is unknown by combining the external data set 35A and the internal data set 35B using the data fusion technology without a common variable. As a result, it is possible to appropriately deliver the advertisement content to the customer of the external data set 35A whose advertisement ID is unknown.

[0204] <Fifth Embodiment>

In a delivery system 30 of the fifth embodiment, a processor 31 is configured to perform delivery control processing illustrated in FIG. 13 instead of the delivery control processing illustrated in FIG. 12. Hereinafter, details of the delivery control processing performed by the processor 31 will be selectively described as a description of the fifth embodiment. The configuration of the delivery system 30 not mentioned in the present embodiment may be understood to be the same as that of the fourth embodiment.

- [0205] In the present embodiment, when a delivery request is input from a user company side system 40 through a communication interface 39, the processor 31 performs the delivery control processing illustrated in FIG. 13.
- [0206] When the delivery control processing is started, the processor 31 acquires, from the user company side system 40, delivery designation information together with advertisement content to be delivered, and a customer data set as the external data set 35A (S710).
- [0207] The delivery designation information acquired in S710 does not include the target designation information but includes only the information on designation of number of deliveries. The customer data set acquired as the external data set 35A is a specific customer data set describing the characteristics of a customer group corresponding to a delivery target narrowed down by the user company.
- [0208] Thereafter, the processor 31 combines the external data set 35A and the internal data set 35B to generate the extended data set 35C, similarly to the processing in S620 (S720). The extended data set 35C includes, for each entity, an extended data in which a feature data included in the external data set 35A of a corresponding customer and a feature data included in the internal data set 35B of a corresponding user are combined.
- [0209] In the processing of S720 of the present embodiment, a result in which all the users of the internal data set 35B each is associated with a customer of the external data set 35A does not occur. The extended data set 35C of the present embodiment also includes, as the extended data of one entity, the feature data of the user not associated with the customer on the use company side. This extended data is the feature data of the user included in the internal data set 35B that is not substantially extended.
- [0210] In the present embodiment, among a group of entities corresponding to the extended data set 35C, a group of entities associated with a customer group corresponding to the external data set 35A is expressed as a seed, and a group of other entities is expressed as a non-seed.

- [0211] After the processing of S720, the processor 31 calculates the similarity of the features indicated by the internal data set 35B between each non-seed entity and each seed entity based on the extended data set 35C (S730). The similarity may be calculated by a distance on the feature space between each non-seeded entity and each seeded entity.
- [0212] After the similarity is calculated, the processor 31 determines, as delivery destinations, entities corresponding to the number of deliveries designated by the delivery designation information in descending order of similarity (S740). At this time, all the entities corresponding to the seed are also determined as the delivery destinations.
- [0213] In this way, the processor 31 selects, as delivery destinations of the advertisement content, the seed which is a set of users associated with customers corresponding to the external data set 35A, and a set of users having similar features to the seed among users corresponding to the internal data set 35B.
- [0214] Thereafter, similarly to the processing of S650, the processor 31 transmits, to the information terminals of the delivery destinations determined in S740, the advertisement content provided from the user company side system 40 through the wide area network (S750). Thereafter, the delivery control processing is ended.
- [0215] According to the delivery system 30 of the present embodiment described above, on the basis of the data set of the customer group provided from the user company side system 40, it is possible to deliver advertisement content to the information terminals of a larger set of consumers exhibiting features similar to the customer group. Therefore, according to the present embodiment, it is possible to efficiently deliver advertisements to many consumers.
- [0216] <Sixth Embodiment>

A delivery system 30 of the sixth embodiment is configured to provide a prediction service together with an advertisement deliver service similar to the delivery system 30 of the fourth embodiment or the fifth embodiment.

[0217] In the present embodiment, a processor 31 performs prediction processing illustrated in FIG. 14 in response to an execution request from a user company side system 40. Hereinafter, details of the prediction processing performed by the processor 31 will be selectively described as a description of the sixth embodiment. The configuration of the delivery system 30 not mentioned in the present embodiment may be understood to be the same as that of the fourth embodiment or the fifth embodiment.

[0218] When starting the prediction processing, the processor 31 acquires a data set to be analyzed from the user company side system 40 together with analysis condition designation information through a communication interface 39 (S810). The data set to be analyzed is a data set including a feature data for each customer to be analyzed.

[0219] The analysis condition designation information may be information designating a product for which the purchase possibility of the customer is evaluated. In the prediction processing, the possibility that each customer to be analyzed purchases a specified target product is predicted by calculating a prediction value of the purchase quantity of the target product. The prediction here corresponds to estimating the behavior of the customer, and the prediction value corresponds to an estimation value regarding the behavior.

[0220] After performing the processing of S810, the processor 31 performs processing similar to the processing of S110 to S170 or S310 to S370 in the analysis processing by using the data set to be analyzed as the first data set 15A and further using the internal data set 35B as the second data set 15B, thereby calculating a correspondence matrix Ω^* indicating a correspondence between each customer to be analyzed and each user having the feature data in the internal data set 35B (S820).

[0221] The processor 31 further extracts, for each customer to be analyzed, a predetermined number of users close to a corresponding customer based on the calculated correspondence matrix Ω^* , and calculates a prediction value of the number of target products purchased by the corresponding customer by a weighted average of the number of target products purchased by the extracted users that can be specified from the internal data set 35B (S830). In this way, the processor 31 estimates the purchasing behavior of the customer from the associated purchasing behavior of the user. The internal data set 35B includes information for which the number of target products purchased by each user can be specified.

[0222] Each element of the correspondence matrix Ω^* indicates the similarity between the customer and the user with a value of 0 to 1. The element in the i-th row and the j-th column in the correspondence matrix Ω^* indicates, by a value 0 to 1, the similarity between the i-th user in the set of users corresponding to the internal data set 35B and the j-th customer in the set of customers corresponding to the data set to be analyzed.

[0223] The weighted average is calculated using, for example, similarity as a weight. Assuming that the first, second, and third users are extracted as three users close to the customer, the weighted average can be calculated as follows.

[0224] That is, when the similarity between the customer and the first user is w_1 , the similarity between the customer and the second user is w_2 , the similarity between the customer and the third user is w_3 , the number of target products purchased by the first user is p_1 , the number of target products purchased by the second user is p_2 , and the number of target products purchased by the third user is p_3 , the prediction value p_e of the number of target products purchased by the customer may be calculated by $p_e = (w_1 \times p_1 + w_2 \times p_2 + w_3 \times p_3)/3$.

[0225] From the correspondence matrix Ω^* , it is possible to specify, for all the users, the similarity between each user and each customer (in other words, the

magnitude of association). Therefore, the prediction value of the number of target products purchased by the customer may be calculated by the weighted average of the number of target products purchased by all the users without the process of extracting the close users.

[0226] After performing the processing of S830, the processor 31 outputs prediction data in which a prediction value of the number of corresponding products purchased for each customer is described to an execution request source of the prediction processing (S840). Thereafter, the processor 31 ends the prediction processing illustrated in FIG. 14.

[0227] According to another example, after performing the processing of S830, instead of or in addition to outputting the prediction data, the processor 31 may perform a processing of delivering the advertisement content for promoting the purchase of the target product to the number of customers corresponding to the number of deliveries designated by the user company in descending order of the prediction value, based on the prediction value of the number of corresponding products purchased for each customer (S840).

[0228] Although the delivery system 30 of the sixth embodiment has been described above, according to the present embodiment, it is possible to provide a meaningful advertisement delivery service and further provide a meaningful marketing solution by using the data fusion technology without a common variable.

[0229] [Others]

It goes without saying that the present disclosure is not limited to the above-described embodiments, and various modes can be adopted. Functions of one element in the aforementioned embodiments may be distributed to two or more elements. Functions of two or more elements in the aforementioned embodiments may be integrated into one element. A part of the configurations of the aforementioned embodiments may be omitted. At least a part of one configuration of the aforementioned embodiments may be added to or replaced with another configuration of the aforementioned

embodiments. All aspects included in the technical ideas identified by the languages recited in the claims are embodiments of the present disclosure.

CLAIMS

1. An information processing system comprising:

a first acquisition unit configured to acquire a first data set relating to first entities, the first data set describing a feature of each of the first entities;

a second acquisition unit configured to acquire a second data set relating to second entities, the second data set describing a feature of each of the second entities;

a dimension reduction unit configured to perform dimension reduction processing on a group of first feature vectors identified from the first data set, each of the first feature vectors representing a feature of a corresponding one of the first entities, and a group of second feature vectors identified from the second data set, each of the second feature vectors representing a feature of a corresponding one of the second entities, to generate a group of first low-dimensional feature vectors corresponding to the group of the first feature vectors and a group of second low-dimensional feature vectors corresponding to the group of second feature vectors, the group of the second low-dimensional feature vectors having the same number of dimensions as the group of the first low-dimensional feature vectors; and

an association unit configured to associate each of the first entities with at least one of the second entities based on the group of the first low-dimensional feature vectors and the group of the second low-dimensional feature vectors.

2. The information processing system according to claim 1, wherein the association unit associates each of the first entities with at least one of the second entities based on similarity between the first entities specified from the group of the first low-dimensional feature vectors and similarity between the second entities specified from the group of the second low-dimensional feature vectors such that mutual relationship between the first entities regarding the similarity matches mutual relationship between the second entities.

3. The information processing system according to claim 1 or 2, wherein:
the group of the first low-dimensional feature vectors is defined by a first feature space;

the group of the second low-dimensional feature vectors is defined by a second feature space; and

the association unit searches for a mapping for mapping the first entities on the first feature space to the second feature space such that distribution of the first entities in the first feature space specified from the group of the first low-dimensional feature vectors matches distribution of the second entities in the second feature space specified from the group of the second low-dimensional feature vectors, and associates each of the first entities with at least one of the second entities based on the mapping.

4. The information processing system according to claim 1 or 2, wherein:
the association unit searches, as a matrix Ω^* , a matrix Ω that maximizes a value $Z(\Omega)$ according to the following formula including a matrix K , a matrix L , and a matrix H ,

[Mathematical formula 1]

$$\begin{aligned}\Omega^* &= \operatorname{argmax}_{\Omega} Z(\Omega) \\ Z(\Omega) &= \operatorname{trace}(K' \Omega^T L' \Omega) \\ K' &= HKH \\ L' &= HLH\end{aligned}$$

and is configured to associate each of the first entities with at least one of the second entities based on the matrix Ω^* ;

the number of the first entities is N ;

the number of the second entities is the same as that of the first entities;

the matrix K is a first similarity matrix with N rows and N columns in which a value of an element in an i -th row and a j -th column represents similarity between an i -th entity and a j -th entity among the first entities, and is calculated based on a first low-dimensional feature vector of the i -th entity among the first entities and a first low-dimensional feature vector of the j -th entity among the first entities;

the matrix L is a second similarity matrix with N rows and N columns in which a value of an element in an i-th row and a j-th column represents similarity between an i-th entity and a j-th entity among the second entities, and is calculated based on a second low-dimensional feature vector of the i-th entity among the second entities and a second low-dimensional feature vector of the j-th entity among the second entities; and

the matrix H is a matrix with N rows and N columns in which a value of an element in an i-th row and a j-th column indicates a value of $1-1/N$ when $i = j$ and indicates a value of 0 when $i \neq j$.

5. The information processing system according to claim 4, wherein

the association unit is configured to improve the matrix Ω^* by repeatedly performing a re-search processing for the matrix Ω^* until a predetermined condition is satisfied, and associate each of the first entities with at least one of the second entities based on the improved matrix Ω^* , and

the re-search processing includes:

changing a dimension reduction method in the dimension reduction processing based on the matrix Ω^* ; and

causing the dimension reduction unit to perform the dimension reduction processing based on the dimension reduction method changed, and

re-searching the matrix Ω^* based on the group of the first low-dimensional feature vectors and the group of the second low-dimensional feature vectors newly obtained by the dimension reduction processing.

6. The information processing system according to claim 5, wherein the association unit changes the dimension reduction method such that a distance in a feature space between a first low-dimensional feature vector and a second low-dimensional feature vector corresponding to each other in the group of the first low-dimensional feature vectors and the group of the second low-dimensional feature vectors is shortened.

7. The information processing system according to any one of claims 1 to 6, wherein:

the first data set includes two or more first feature data, and each of the two or more first feature data represents a feature of a corresponding one of the first entities;

the second data set includes two or more second feature data, and each of the two or more second feature data represents a feature of a corresponding one of the the second entities; and

the information processing system further comprises

a data fusion unit that generates an extended data set including two or more extended data, each of the two or more extended data being a combined data of one of the two or more first feature data and a corresponding one of the two or more second feature data, by combining each of the two or more first feature data with one of the two or more second feature data based on association between the first entities and the second entities by the association unit.

8. The information processing system according to any one of claims 1 to 7, wherein the first entities and the second entities are persons,

the first data set is a data set describing a first feature of each person belonging to a first group, and

the second data set is a data set describing a second feature of each person belonging to a second group.

9. The information processing system according to claim 8, wherein a combination of the first feature and the second feature is a combination of a feature related to purchasing behavior, and a feature related to movement in at least one of an online space and an offline space and/or a feature related to visits to two or more points in at least one of the online space and the offline space.

10. The information processing system according to any one of claims 1 to 9, wherein

the first entities and the second entities are persons,

identification information of an information terminal corresponding to each of the second entities is associated with the second data set, and

the information processing system further comprises:

a selection unit that selects, as delivery destinations of information content, at least a part of a set of second entities associated with any one of the first entities by the association unit; and

a delivery unit configured to deliver the information content to a set of information terminals corresponding to the delivery destinations of the information content based on the identification information.

11. The information processing system according to claim 10, wherein the selection unit selects, as delivery destinations of the information content, a first set that is the set of second entities associated by the association unit with any one of the first entities, and a second set having similar features to the first set among the second entities.

12. The information processing system according to any one of claims 1 to 11, wherein the first entities and the second entities are persons,

the second data set describes a feature related to behavior of each of the second entities,

the information processing system further includes

an estimation unit that calculates, for each of one or more entities of interest, an estimation value regarding behavior of a corresponding entity of interest,

the one or more entities of interest are at least some of the first entities,

and

the estimation value is calculated based on a feature related to the behavior of at least one of the second entities associated with the corresponding entity of interest.

13. A computer-implemented information processing method comprising:
 - acquiring a first data set relating to first entities, the first data set describing a feature of each of the first entities;
 - acquiring a second data set relating to second entities, the second data set describing a feature of each of the second entities;
 - performing dimension reduction processing on a group of first feature vectors identified from the first data set, each of the first feature vectors representing a feature of a corresponding one of the first entities, and a group of second feature vectors identified from the second data set, each of the second feature vectors representing a feature of a corresponding one of the second entities, to generate a group of first low-dimensional feature vectors corresponding to the group of the first feature vectors and a group of second low-dimensional feature vectors corresponding to the group of second feature vectors, the group of the second low-dimensional feature vectors having the same number of dimensions as the group of the first low-dimensional feature vectors; and
 - associating each of the first entities with at least one of the second entities based on the group of the first low-dimensional feature vectors and the group of the second low-dimensional feature vectors.

14. The computer-implemented information processing method according to claim 13, wherein

- the association includes
 - associating each of the first entities with one of the second entities based on similarity between the first entities specified from the group of the first low-dimensional feature vectors and similarity between the second entities specified from the group of the second low-dimensional feature vectors such that mutual relationship

between the first entities regarding the similarity matches mutual relationship between the second entities.

15. A computer-readable recording medium storing a computer program including instructions for causing a computer to perform the information processing method according to claim 13 or 14.