Big Five Allocation Method to Household Members in Synthetic Population

Yugo Kitajima[1] and Tadahiko Murata[1]

***Abstract*— Synthetic population data is virtual household data synthesized from statistics that are publicly released. It includes synthesized attributes such as age, gender, employment status of household members, and other details that are utilized for social simulations for target areas or communities. It has basic attributes of household members, though, it does not have psychological attributes that strongly relate to individual decision-making. As prior research to assign psychological attributes to household members, a method has been proposed to assign the Big Five personality traits to each household member in synthetic population data. However, since this method assigns the Big Five based solely on gender from the reference survey, it may fail to consider effect of ages in their decision making. Therefore, in this paper, we propose a Big Five assignment method that takes gender and age into account and quantitatively compared its reproducibility and validity. Furthermore, using external data, we analyzed the relationship between lifestyle indicators corresponding to the Big Five and areas of interest related to behavioral choices, and presented clues for designing decision-making models using the Big Five.**

I. INTRODUCTION

In recent years, social simulation has become increasingly important in fields such as social issue analysis and policy formulation. In social simulation, agent decision-making has a significant impact on the final outcome. Therefore, even if attributes such as age and gender are identical, behavior may differ due to variations in psychological attributes. Simulations that do not account for such individual differences face the challenge of being unable to adequately represent the behavioral structures of real society.

Against this backdrop, Murata and Uetani [1] conducted a prior study in which they assigned the Big Five psychological traits to synthetic population data. The Big Five consists of five personality traits: Extraversion (E), Agreeableness (A), Conscientiousness (C), Neuroticism (N), Openness (O) [2]. Existing research has shown that these traits exhibit not only gender differences but also age differences [3]. Consequently, as in previous studies, if age is not taken into account in Big Five assignment methods for synthetic population data, the Big Five may not be assigned appropriately, and social simulations may fail to sufficiently reproduce the diversity of behavior observed in real society.

*This work was supported by JST Mirai Program (JPMJMI23B1). The synthetic population data used in this study were supported by the Joint Usage/Research Center for Interdisciplinary Large-Scale Information Infrastructures (jh260056) and the High Performance Computing Infrastructure (hp240468, hp250520).

[1]Yugo Kitajima and Tadahiko Murata are with the D3 Center, The University of Osaka, Osaka, Japan. tadahiko.murata.cmc@osaka-u.ac.jp

Therefore, in this study, we propose a Big Five assignment method based on age groups and gender and compare it with multiple assignment models, including existing methods. In the evaluation, we verify the consistency and stability of the distribution not only in terms of overall agreement with the Big Five distribution estimated from the reference survey but also from the perspective of the degree of agreement in the distribution by age group. Furthermore, since the relationship between the Big Five and behavioral choices is important for connecting to decision-making models in social simulations, we conduct a basic verification using external data regarding the correlations between the Big Five and behavior used in this study.

II. RESEARCH METHOD

A. Models for Comparison

In this section, we explain five models to assign Big Five factors in this study. We analyze the models from two aspects: (1) the estimation of probability distributions for the Big Five based on reference survey data, and (2) the assignment of the Big Five to synthetic population data. This overview clarifies which conditional variables each model takes into account and which distribution structures (marginal distributions, joint distributions, and correlation structures) each model aims to reproduce.

In this study, we use the "Japan Household Panel Survey on Consumer Preferences and Satisfaction" [4] conducted by the Institute of Social and Economic Research at The University of Osaka as the reference survey for the Big Five. In accordance with the design of the reference survey, the target age range is limited to 20–70 years. Furthermore, we use scores calculated based on the TIPI [5] for the Big Five.

In the following, based on the distributions and statistics obtained from the reference survey, we describe in detail, from the perspective of probability distributions and sampling, a model for assigning the Big Five to synthetic population data. We compare the following four models to allocate psychological factors to individuals in synthetic population with age 20 to 70. In the four models employed in this study, we extract the value of psychological factors according to the following empirical distributions:

Model 0: Gender-conditional marginal distributions,
Model 1: Gender-conditional joint distribution (Previous),
Model 2: Gender $\times$ age group-conditional joint distribution (Proposed),
Model 2h: Gender $\times$ age group $\times$ household-conditional joint distribution,

Model 3: Gaussian Copulas for Reproducing Correlation Structures.

Before describing the detail of the above four models, we define the notation used throughout this study.

B. Common Notation

[Individual Attributes]

Each individual in the reference survey and in the synthetic population data has attributes such as gender and age. Let gender be $s \in 1, 2$ where $(\text{male} = 1, \text{female} = 2)$, and age group be $a \in A$ where the set of age groups A is described as follows:

$$A = \{20\text{–}29, 30\text{–}39, 40\text{–}49, 50\text{–}59, 60\text{–}69, 70\}. \quad (1)$$

Ten-year age groups are employed except Age 70 in this study since the number of samples in each age group significantly decreases when one-year age groups are used. H is a set of five categories of household type, defined as,

$$H = \begin{Bmatrix} \text{Single}, \\ \text{Couple}, \\ \text{Couple and children}, \\ \text{A parent and children}, \\ \text{Others}. \end{Bmatrix} \quad (2)$$

Let $i = 1, 2, ..., N$ be IDs for individuals in the synthetic population data, and the sex, age group, and household composition of individual i are denoted by s_i, a_i, and h_i, respectively. For example, a_i = 30–39 indicates that individual i belongs to the age group corresponding to "individual i is in their 30s."

[Discrete Representation of the Big Five]

The five factors of the Big Five (E: Extraversion, A: Agreeableness, C: Conscientiousness, N: Neuroticism, O: Openness) are combined

$$\mathbf{x} = (x_E, x_A, x_C, x_N, x_O) \in \{1, ..., 7\}^5, \quad (3)$$

where $\mathbf{x}$ is a vector consisting of the combined scores for the five factors, where each factor takes discrete values from 1 to 7. For example, $\mathbf{x} = (5, 4, 6, 2, 3)$ represents the Big Five score set for a single individual, where x_E = 5, x_A = 4, x_C = 6, x_N = 2, and x_O = 3. Consequently, there are 7^5 (= 16,807) possible combinations of Big Five scores. Let the Big Five vector assigned to individual i in the synthetic population data be denoted by $\mathbf{x}_i$, and let the component corresponding to factor $t \in \{E, A, C, N, O\}$ be denoted by $x_{i,t}$ (e.g., $x_{i,E}$ is the extraversion score of individual i).

[Continuous Representation of the Big Five]

To estimate the correlation matrix among factors t, we also use continuous scores derived from TIPI. We denote the continuous scores as follows:

$$\mathbf{z} = (z_E, z_A, z_C, z_N, z_O) \in R^5, \quad (4)$$

where $\mathbf{z}$ is the continuous value derived from TIPI after inversion and averaging, and we use this $\mathbf{z}$ to estimate the correlation matrix. On the other hand, the discrete scores $\mathbf{x}$ from 1–7 are used to construct the marginal distributions.

[Probability Distribution and Empirical Distribution]

The probability distribution estimated from the reference survey is denoted as $P(\cdot)$. Note that the $P(\cdot)$ used in this study is an empirical distribution which is obtained by normalizing the frequencies observed in the reference survey data as probabilities. For example, in a population of gender s, if $\mathbf{x}$ is observed $n(\mathbf{x}, s)$ times and the total number is $n(s)$, the probability distribution is defined as follows:

$$P(\mathbf{x}|s) = n(\mathbf{x}, s)/n(s). \quad (5)$$

The marginal distribution $P(t = k|s)$ is similarly obtained by counting the number of times when factor t takes the value k and normalizing it.

[Sampling]

Psychological factor t is allocated as extracting random numbers from a probability distribution. Using a symbol $\sim$, a variable $\mathbf{x}_i$ is once generated according to the probability distribution $P(\cdot)$ as follows:

$$\mathbf{x}_i \sim P(\mathbf{x}|\cdot). \quad (6)$$

Through this operation, the allocated psychological factors in synthetic population data set are expected to reproduce the statistical characteristics of the reference distribution.

C. M0: Gender-Conditional Marginal Distributions

M0 is the simplest baseline model, assuming that each psychological factor of the Big Five factors is independent one another and follows only the marginal distributions by gender. By intentionally ignoring correlations between psychological factors (e.g., the tendency for E and O to be high simultaneously), it serves as a baseline for comparison with models that account for correlation structures.

Distribution Estimation: The reference survey is divided by gender s, and the empirical distribution of each psychological factor $t \in E, A, C, N, O$ is described as follows:

$$P(t = k|s), k \in \{1, 2, ..., 7\}. \quad (7)$$

$P(t = k|s)$ represents the "probability that factor t takes score k in the population of gender s."

Allocation: For each individual i in the synthetic population data, extract factor t independently according to the empirical distribution of gender s_i as follows:

$$x_{i,t} \sim P(t|s_i), t \in \{E, A, C, N, O\}. \quad (8)$$

In this case, the joint distribution conditional on gender simply becomes

$$P(x|s) = \prod_{t \in \{E,A,C,N,O\}} P(t|s). \quad (9)$$

Therefore, correlations between factors in the reference distribution are not kept.

D. M1: Allocation Based on Gender-Conditional Joint Distribution (Previous Method)

M1 is a method used in previous research that achieves allocation while preserving correlations between factors by utilizing the joint distribution of the Big Five for each gender.

Distribution Estimation: From the reference survey, estimate the following distribution as the empirical distribution for each gender s.

$$P(\mathbf{x}|s). \tag{10}$$

This represents the "probability of observing a set of five factors x in the population of gender s." Specifically, count the frequency with which the same $\mathbf{x}$ is observed and normalize it by dividing by the total number.

Allocation: For individual i in the synthetic population data, the values of five factors are simultaneously extracted using the empirical distribution $P(\mathbf{x}|s_i)$ for the gender of the individual s as follows:

$$\mathbf{x}_i \sim P(\mathbf{x}|s_i). \tag{11}$$

This operation reflects the gender differences and inter-factor correlations of the Big Five.

E. M2: Allocation Based on Gender × Age Group Conditional Joint Distribution (Proposed Method)

M2 is the proposed method in this study. By introducing age group as a conditional variable in addition to gender, it aims to reproduce the age-dependent distribution structure (mean, variance, and combinations) of the Big Five.

Distribution Estimation: Stratify the reference survey by gender s and age group a, estimate the following distribution as the empirical distribution for each gender s and age group a.

$$P(\mathbf{x}|s, a). \tag{12}$$

$P(\mathbf{x}|s, a)$ represents the "probability of observing the set $\mathbf{x}$ in the population of gender s and age group a."

Allocation: For each individual i in the synthetic population data, the values of five factors are simultaneously extracted using the empirical distribution $P(\mathbf{x}|s_i, a_i)$ for the gender and the age group of the individual i as follows:

$$\mathbf{x}_i \sim P(\mathbf{x}|s_i, a_i). \tag{13}$$

It should be noted that stratifying by (gender × age group) may result in a small sample size within each cell since the empirical distribution depends on the sample size of the reference survey. In this case, the variance of the estimated $P(\mathbf{x}|s, a)$ may increase. Therefore, in the next section, we verify the results by combining cell-level goodness-of-fit and stability indices.

F. M2h: Allocation Based on Gender × Age Group × Household-Conditional Joint Distribution

M2h is an extension of M2 that attempts a more detailed Big Five allocation by introducing household composition h as a conditional variable in addition to gender and age group. Since household composition is an attribute that may be related to social behavior and lifestyle, it is thought to potentially influence the distribution of psychological traits.

On the other hand, the addition of conditional variables significantly subdivides the cells, making it easy for the reference survey to suffer from insufficient sample sizes within cells. In this study, we regard M2h as a comparison model to verify model behavior including the estimation instability caused by such subdivision.

Distribution Estimation: From the reference survey, estimate the following distribution as the empirical distribution for each gender s, age group a, and household composition h.

$$P(\mathbf{x}|s, a, h). \tag{14}$$

Here, $P(\mathbf{x}|s, a, h)$ represents "the probability that $\mathbf{x}$ is observed in a population satisfying gender s, age group a, and household composition h."

Assignment: For individual i in the synthetic population data, the values of five factors are simultaneously extracted using the empirical distribution $P(\mathbf{x}|s_i, a_i, h_i)$ for the gender, the age group and the household composition of the individual i as follows:

$$\mathbf{x}_i \sim P(\mathbf{x}|s_i, a_i, h_i). \tag{15}$$

G. M3: Gaussian Copulas to Reproduce Correlation Structures

M3 is a method that models the "marginal distributions of each factor" and the "correlation structure between factors" separately, rather than directly maintaining the joint distribution $P(\mathbf{x}|s)$ as the empirical distribution. That is, in M3, correlations are estimated from continuous scores, and marginal distributions are estimated from discrete scores.

This type of separative approach has been widely used in the fields of financial engineering and risk management to construct multivariate distributions that preserve the correlation structure from limited samples. In this study, we do not intend to establish a direct correspondence with Big Five allocation methods in social simulation. Rather, we introduce M3 as a point of comparison with "general methods that explicitly preserve the correlation structure."

Distribution estimation: For each gender s, the correlation matrix R_s is estimated from the reference survey. Here, R_s is a 5 × 5 matrix and is calculated using the continuous scores $\mathbf{z} = (z_E, z_A, z_C, z_N, z_O)$ as

$$R_s = \mathrm{Corr}(\mathbf{z}|s). \tag{16}$$

On the other hand, the marginal distributions are estimated using discrete scores $\mathbf{x}$, and the cumulative distribution functions (CDFs) $F_{t,s}$ are constructed from the discrete marginal

distributions $P(t = k|s)$ for each factor t. The CDFs are given by

$$F_{t,s}(k) = P(t \leq k|s), (k = 1, ..., 7), \tag{17}$$

and represents the "probability that factor t is less than or equal to k in the population of gender s."

Assignment: For individual i of gender s_i, perform the following. First, generate

$$y \sim N(\mathbf{0}, R_{s_i}) \tag{18}$$

where, $y = (y_E, y_A, y_C, y_N, y_O)$ is a 5-dimensional normal random vector with covariance (correlation) structure R_{s_i}. Furthermore, $N(\mathbf{0}, R_{s_i})$ represents a 5-dimensional multivariate normal distribution with mean vector $\mathbf{0}$ and correlation matrix R_{s_i}. Next, using the cumulative distribution function (CDF) ϕ of the standard normal distribution, we transform it as

$$u_t = \phi(y_t) \tag{19}$$

At this point, $u_t \in (0, 1)$ becomes a random variable following a uniform distribution (probability integral transformation).

Finally, map 'u_t' to a "discrete score between 1 and 7 following the reference marginal distribution." For discrete distributions, define the inverse CDF (quantization function) as

$$F^{-1}_{t,s}(u) = \min\{k \in 1, ..., 7|F_{(t,s)}(k) \geq u\} \tag{20}$$

and obtain the score via

$$x_{i,t} = F^{-1}_{t,s_i}(u_t) \tag{21}$$

to obtain the score. This operation means "transforming the uniform random variable u_t into a discrete value following the marginal distribution $P(t|s)$ of the reference survey."

It should be noted that M3 matches the marginal distribution $P(t|s)$ to the reference survey and reproduces the correlation structure under the Gaussian copula assumption. However, it is important to note that, unlike M1, which directly retains the actual joint distribution $P(\mathbf{x}|s)$ as the empirical distribution, higher-order features of the joint distribution (e.g., the concentration of specific combinations in the Big Five) are approximated by the copula assumption.

III. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we compare the results of assigning the Big Five traits to synthetic population data [6] using the Big Five assignment models (M0, M1, M2, M2h, M3) with those of the reference survey [4], evaluating: (1) the consistency of the overall distribution, (2) the consistency of the stratified structure by age group (and by gender $\times$ age group), (3) the reproducibility of the correlation structure among the Big Five factors. For the evaluation, since the reference survey was limited to the 20–70 age group, the range of the synthetic population data was also fixed at 20–70 years. Here, Sumiyoshi Ward, Osaka City, Osaka Prefecture, Japan (Population was 154,239 in 2015 National Census) is selected for the Big Five assignment. First, let us describe the evaluation metrics we employ in this paper.

A. Evaluation Metrics

In this study, we evaluate from multiple perspectives the extent to which the five-dimensional personality trait distribution comprising the Big Five (E, A, C, N, O) aligns between the reference survey and the synthetic population data. Each metric is used to quantify different aspects: the overall consistency of the distribution, the reproducibility of the inter-trait structure, and the reproducibility of the age-group structure.

[Consistency of the Joint Distribution]

To evaluate the consistency of the joint distribution (E, A, C, N, O) that simultaneously considers the five Big Five factors, we use the Jensen–Shannon distance (JS distance). JS distance is a symmetric and bounded distance obtained by averaging the Kullback-Leibler distance—which has an asymmetric characteristic depending on which of the two distributions is used for measurement—after measuring the distance in both directions. It is numerically stable for discrete distributions and has characteristics that are easy to interpret.

[Reproducibility of Inter-Factor Structures]

To evaluate whether the relational structure among the five Big Five factors is reproduced, we define $corr$ as an index based on Kendall's rank correlation for the off-diagonal elements of the inter-factor correlation matrix. Example: Calculate the Kendall correlation matrix for the five factors in both the reference and model distributions, and define $corr$ as the RMSE of the 10 ($= {}_5C_2$) off-diagonal elements. Here, RMSE refers to the root mean square error. This metric quantitatively evaluates "the extent to which the strength and direction of the relationships between each factor correspond to the reference survey," capturing a structural aspect distinct from the fit of the joint distribution (JS distance). Here, this metric is the RMSE of the off-diagonal elements of the inter-factor correlation matrix. The closer it is to 0, the closer it is to the reference.

[Reproducibility of the Age Group Structure]

To evaluate the "structural reproducibility by age group" that is the objective of this study, we assess the differences in the mean and variance of each factor for each gender $\times$ age group cell using RMSE. In the table, the error regarding the mean is denoted as "mean," and the error regarding the variance is denoted as "var."

These indicators evaluate the "reproducibility of the Big Five levels and variability when viewed by age group," and the evaluation target is clearly different from the overall joint distribution match (JS distance).

[Distribution Diversity]

To supplement the assessment of whether the shapes of the distributions are "similar" by also determining "how diverse" they are, the difference in joint entropy is calculated using the following formula.

TABLE I
EVALUATION RESULTS IN JAPAN)

Model	JS	corr	mean	var	Δ
M0	0.18935	0.154	0.192	0.185	0.839
M1	0.00067	0.001	0.215	0.186	-0.007
M2	0.01671	0.008	0.008	0.014	-0.030
M2h	0.05411	0.022	0.149	0.209	-0.236
M3	0.15747	0.043	0.213	0.186	0.672

$$\Delta H = H_{\text{model}} - H_{\text{ref}} \tag{22}$$

A value of $\Delta H > 0$ indicates that the model distribution has greater diversity than the reference distribution, while $\Delta H < 0$ indicates that the model distribution has less diversity than the reference distribution. However, it is important to note that entropy is an indicator that summarizes the overall shape of a distribution and does not directly evaluate individual age classes or specific combination structures.

B. Results of Allocation

[Consistency with the Overall Distribution]

Table I shows the key metrics evaluated against the reference distribution, which includes all ages and genders. The best values for each evaluation metric are underlined. While M1 was the best for overall consistency (JS) and correlation reproduction (corr), M2 was the best for age group structure (mean, var). This indicates that the model considered optimal may vary depending on the evaluation objective (whether emphasis is placed on the consistency of the overall aggregate or on the consistency of the stratified structure). Since the objective of this study is to reproduce the stratified structure, we do not judge the superiority of models based solely on the reference distribution that includes all ages, but rather interpret the results by including evaluations by age group.

This result may be due to the fact that the main indicators evaluated against the reference distribution, which includes all ages and genders, are recalculated for the "distribution combining age-class distributions," which can average out discrepancies in the stratified structure. In particular, for nonlinear distances such as the JS distance, the "weighted average of stratified distances" and the "distance to the combined distribution" generally do not match.

[Consistency of Age-Group Distributions]

Next, we examine the degree of agreement by age group. Fig. 1 shows the the JS distance by age group when evaluated against a reference distribution that includes all ages and genders, and M2 exhibits a relatively small (i.e., better) JS distance in all age groups. In this section, we present the JS distance as a representative metric because it is easy to interpret as a symmetric distance between probability distributions and possesses the property of allowing stable comparisons even for discrete distributions. Furthermore, as shown in Fig. 2, which illustrates the "variation in JS

Fig. 1. JS Distance by Age Group

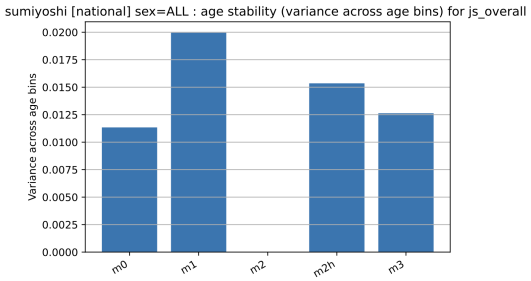

Fig. 2. JS Distance by Age Group

distance across age groups," M2 exhibits the smallest value (approximately 1.12×10^{-6}), suggesting that the bias—where performance deteriorates only in specific age groups—is relatively small.

On the other hand, because M2h adds household composition as a condition, the number of samples in the reference cells decreases, and as a result, the indicators by age group tend to deteriorate compared to M2. However, this deterioration does not immediately imply that "adding the household composition condition is inherently inappropriate." Since M2h subdivides the condition space, noise increases in empirical distribution estimation, and as a result, the JS distance may deteriorate. In other words, it is natural to interpret the behavior of M2h as a trade-off between "refinement of allocation through the addition of conditions" and "stability of estimation due to sample size constraints," and the deterioration of the evaluation metrics can be considered a manifestation of estimation instability.

In addition, as an example of evaluation by age group, Table II shows the evaluation results limited to the 40–49 age group (gender = all). In this age group, M2 showed relatively good values for the JS, corr, and mean/var metrics, and this was also the case for other age groups.

C. Big Five

In this section, to verify that the Big Five assignment results for the synthetic population data do not significantly contradict external data, we compare the lifestyle response rates from the National Behavioral Survey 8) with the proportion of high Big Five scores (scores of 6–7). However, since this comparison involves matching data sets with different

TABLE II

EVALUATION RESULTS IN AGE 40-49

Model	JS	corr	mean	var	Δ
M0	0.308	0.196	0.105	0.051	1.427
M1	0.173	0.022	0.137	0.054	0.598
M2	0.001	0.005	0.004	0.009	-0.004
M2h	0.154	0.023	0.055	0.122	0.304
M3	0.277	0.061	0.134	0.052	1.252

TABLE III

COMPARISON OF BIG FIVE ALLOCATION RATIOS AND BEHAVIORAL SURVEY RESPONSE RATES

Big 5	Attitude	Allocation	Survey	Δ
E	Sociability	0.153	0.152	0.001
O	Culture	0.100	0.150	0.050

measurement subjects and scales, it should be viewed as a supplementary check to confirm the validity of the results rather than a rigorous validation.

The correspondence between the Big Five and lifestyle is based on the following assumptions:

$$E \leftrightarrow \text{Sociability}, O \leftrightarrow \text{Cultural Orientation}$$

This correspondence is based on descriptions in McCrae and John [2], such as "Extraversion = Sociability/Activity" and"Openness = Intellectance" Since there were no explicitly corresponding lifestyles for the A, C, and N attributes in the Big Five, we verify this using E and O in this study. Table III compares the percentage of high scorers in the Big Five assignment results with the response rates for corresponding lifestyle items in the National Behavior Survey. While the difference was extremely small for the correspondence between Extraversion (E) and "sociable," the difference was slightly larger for the correspondence between Openness (O) and "cultured" compared to E. This difference may reflect differences in the conceptual scope encompassed by each trait.

IV. CONCLUSIONS

The aim of this study was to achieve a Big Five allocation that preserves the stratified structure of gender $\times$ age group as a foundation for representing heterogeneity in decision-making and behavior in real-scale social simulations. In the evaluation, based on the results regarding the consistency and stability of the distribution—considering not only the overall consistency with the Big Five distribution estimated from the reference survey but also the consistency of the distribution by age group—the proposed method performed best in terms of the reproducibility and stability of the distribution by age group. Based on the above, we judge that it is appropriate to adopt M2 as the Big Five assignment model for synthetic population data at this stage.

Furthermore, to provide supplementary evidence of the significance of assigning the Big Five to synthetic population data, we conducted an analysis using external data. Although this analysis does not directly measure the Big Five, it demonstrated that a consistent correlation structure exists between lifestyle and interests, and that this correlation can change structurally depending on age group and city size.

REFERENCES

[1] T. Murata and M. Uetani, Assignment of Big Five Personality Attributes in Synthetic Populations, *Proceedings of Workshop of Technical Committee on Social Systems, Society of Instrument and Control Engineers*, vol.31, 1 page (2023, in Japanese).

[2] R. R. McCrae and O. P. John, An Introduction to the Five-Factor Model and Its Applications, *Journal of Personality*, vol.60, no.2, pp.175-215 (1992).

[3] C. J. Soto, O. P. John, S. D. Gosling and J. Potter, Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample, *Journal of Personality and Social Psychology*, vol.100, pp.330-348 (2011).

[4] Institute of Social and Economic Research, The University of Osaka, *Japan Household Panel Survey on Consumer Preferences and Satisfaction* (2012).

[5] A. Oshio, S. Abe and Pino Cutrone, Development, Reliability, and Validity of the Japanese Version of Ten Item Personality Inventory (TIPI-J), *the Japanese Journal of Personality*, Vol.21 No.1, pp.40–52 (2012, in Japanese).

[6] T. Murata, T. Harada and D. Masui, Comparing Transition Procedures in Modified Simulated Annealing-Based Synthetic Reconstruction Method without Samples, *SICE Journal of Control, Measurement, and System Integration*, vol.10, no.6, pp.513-519 (2017).