Workplace Assignment to Workers in Synthetic Populations in Japan

Tadahiko Murata, *Senior Member, IEEE*, Daiki Iwase, *Student Member, IEEE*, and Takuya Harada, *Member, IEEE*

***Abstract*— In this article, we assign workplace attributes to each worker in each household in a synthetic population using multiple censuses conducted in Japan. The synthetic population is a set of artificial individual attributes for each resident that is synthesized according to census data. We have synthesized a set of the synthetic populations of Japan. We assign a workplace attribute to each worker to estimate daytime population distribution and develop activity-based models in agent-based or microsimulations. Although statistical information in a residential area or a working place is released by the government and some individual moving data are released by cellphone companies, it is hard to collect the information with home and workplace location of a worker with their family and working information. We employ origin–destination–industry (ODI) statistics to estimate workplaces for workers. Since some attributes in ODI statistics are not available for privacy reasons, we propose a workplace assignment method for all cities, towns, and villages using restricted ODI and OD statistics in Japan. We show how much difference there are between the number of workers using the complete ODI statistics and the number of workers by the proposed workplace assignment method. We show that 88.2% of workers in a city in Japan are assigned to correct cities as workplaces by our proposed method. We also show several maps of daytime population distributions by our proposed method. Synthetic populations with workplace attributes enable real-scale social simulations to design transport or business systems in times of peace or to estimate victims and plan recoveries in times of emergency, such as disasters or pandemics.**

***Index Terms*— Agent-based modeling, daytime population distribution, microsimulations, social simulations, synthetic population, workplace assignment.**

I. Introduction

SYNTHETIC populations attract many researchers who conduct social simulations in a specific geographical area or region. During the COVID-19 pandemic, many researchers tried to see how infected patients spread in a target region [1], [2], [3], [4], [5], [6]. To observe such an increase of patients in a specific region or area using social simulations, researchers need a synthetic population with attributes of each resident and household composition in the target area or region. Currently, there are few research projects, which publicly release the data on country-level synthetic populations only in the U.S. [7], [8], the U.K. [9], Japan [10], [11], [12], and Belgium [13]. As a Japanese team, we distribute the Japanese synthetic population to researchers who are to conduct social simulations for regions in Japan.[1]

The first method to synthesize populations based on statistics, called the synthetic reconstruction method (SR method), was proposed by Wilson and Pownall [14]. They reconstruct individual and household data from statistics with some actual samples using an iterative proportional fitting (IPF) procedure [15]. Barthelemy and Toint [16] indicate that the SR method has difficulty reconstructing populations simultaneously to fit both individuals and households. To cope with this difficulty, Gargiulo et al. [17] and Barthelemy and Toint [16] proposed synthetic population methods without sample data. Lenormand and Deffuant [18] compare the methods with the SR method and show that the synthetic population method without samples can reconstruct a better solution. We employed an SR method without samples in synthesizing Japanese populations [10].

A synthetic population with attributes of residents and households is synthesized based on the census collected from residents in a municipality. Therefore, a synthetic population shows the information on residents as a so-called "night-time population," that is, attributes of individual residents and their households are connected with their home location. In order to estimate the activities of residents in agent-based or microsimulations, it is essential to estimate the movements of residents during the daytime. Since youths and children are going to school mainly by foot or bicycles, we estimate that their schools exist not so far from their homes. On the other hand, locations of workplaces are scattered over areas, including their home and neighboring cities by workers. In this article, we assign workplaces based on workers' statistics available in Japan that helps to estimate their daytime activities.

In order to estimate workplaces for workers, many researchers employ origin–destination (OD) surveys. Workplace assignment methods become different whether their OD

Manuscript received 29 March 2022; revised 9 August 2022; accepted 12 October 2022. This work was supported in part by the Joint Usage and Research Center, Research Institute for Socionetwork Strategies, Kansai University; in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant 20K10362; in part by the Kansai University Fund for Supporting Outlay Research Centers, 2020–2021; in part by the Japan Science and Technology Agency (JST)-Mirai Program under Grant JPMJMI20B3; in part by Joint Usage/Research Center for Interdisciplinary Large-scale Information Infrastructures under Project jh210040-MDH and Project jh220051; and in part by the High Performance Computing Infrastructure in Japan under Project hp210290. *(Corresponding author: Tadahiko Murata.)*

Tadahiko Murata and Daiki Iwase are with the Faculty of Informatics, Kansai University, Takatsuki, Osaka 569-1095, Japan (e-mail: murata@kansai-u.ac.jp).

Takuya Harada is with the College of Systems Engineering and Science, Shibaura Institute of Technology, Saitama 337-8570, Japan (e-mail: t-harada@shibaura-it.ac.jp).

Digital Object Identifier 10.1109/TCSS.2022.3217614

[1] Synthetic Populations of Japan are distributed from the following website: http://www.res.kutc.kansai-u.ac.jp/~murata/rsss-distribution/?lang=en

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

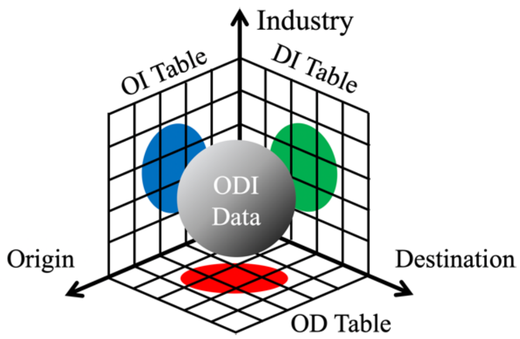


Fig. 1. ODI data and three tables by OD, OI, and DI.

survey is: 1) a sampling survey that extracts several subjects from the entire population or 2) a complete survey that collects replies from the entire population.

Many workplace assignment methods employ 1) sampling OD surveys. For example, Abdel-Aal [19] conducted a workplace assignment using a sampling OD survey among 15 zones in Alexandria, Egypt. The sampling OD survey is adjusted using the IPF procedure to estimate the entire movements between 15 zones in Alexandria.

As an example of 2) a complete OD survey, Fournier et al. [20] employed longitudinal employer household dynamics (LEHD) origin–destination employment statistics (LODES) collected by the Center for Economic Studies of the U.S. Census Bureau[2] and Ye et al. [21] employed Chinese National Population and Economic Census. Fournier et al. [20] utilize three LODES statistics, such as workplace OD totals due to the lack of origin–destination–industry (ODI) data, workplace origin totals by industries (OI), and workplace destinations by industries (DI). Fig. 1 shows the relation among the three attributes (origin, destination, and industries) in the three tables. As shown in Fig. 1, each of the three tables is a projection of the ODI data using two attributes. Since they can utilize only projected tables, they should generate a joint distribution from OI, OD, and DI tables.

In Japan, the complete ODI data (the number of relations among origin–destination–industry) is released by local governments whose area have more than 200 000 residents. However, the destinations category is not fine for cities, towns, or villages, where the number of residents is less than 200 000. In those areas, only six categories are shown for the destination, such as "At home," "In the same city," "In another city in the same prefecture," "In another city in another prefecture," "Unknown in Japan or outside Japan," and "Not available." This article proposes a workplace assignment method for workers in all cities, towns, and villages in Japan (as another unit, there are "wards" in Tokyo and some larger cities). We compare the proposed method with a method using statistics available in a city with more than 200 000 residents to see challenges in the workplace assignment to workers in the proposed method.

This article consists of the following sections. Section II shows a workplace assignment method using the ODI data for all cities, towns, and villages in Japan. We first assign a workplace at the city level (or town and village-level) for each worker. Because restricted ODI data are only available for cities, towns, and villages with less than 200 000 residents, we also employ the OD data with the names of cities, towns, and villages as destinations to assign workplaces. On the other hand, cities with more than 200 000 residents release more precise ODI data. We apply the proposed method with restricted statistics to a city with more than 200 000 residents and then compare the results with the workplace assignment using the complete ODI data for that city. Section III shows a method of workplace assignment at a small area level in an assigned city in Section II. Section IV shows the results of the proposed workplace assignment method for workers. It shows a tendency of workplaces by industries. Section V concludes and shows some further applications of the synthetic population with workplace attributes.

II. Workplace Assignment at City Level

A. Attributes of Household Members in Synthetic Population

Fig. 2 shows attributes for each household member in a synthetic population synthesized by our method [6], [7], [8]. Four address attributes are assigned to each household, such as (A) prefecture, (B) city, (C) small area in the city, and (D) home coordinate (latitude and longitude). In Fig. 2, Japan had more than 127 million residents in 2015. It has 47 prefectures, the biggest one is Tokyo, with more than 13.5 million residents, and the smallest one is Tottori, with 573 thousand. Fig. 2 shows Osaka Prefecture with more than 8.8 million. It has 33 cities, 9 towns, and 1 village. Among them, eight cities have more than 200 000 residents, and the others have less than it. Fig. 2 shows an example of a city with more than 200 000, Takatsuki-city in Osaka. The city has 448 small areas indicated by red lines in the figure. Geospatial Information Authority of Japan, Ministry of Land, Infrastructure, Transport and Tourism, Japan releases the fundamental geospatial data, including the shape data of each building with latitude and longitude. Each household in our synthetic population has a coordinate based on the fundamental geospatial data.

The synthetic population also includes seven biological and social attributes for each household member, such as (1) age, (2) sex, (3) role in the household, (4) working status, (5) working industry, (6) industry size, and (7) income if it is working. The role is assigned to each household member according to its family type, such as husband, wife, child, and parent. We employ nine family types shown in Fig. 3. These nine family types cover 95% of whole households in Japan.

The government of Japan collects census data every five years. Attributes (A) and (B) and (1)–(3) of each household in Fig. 2 are synthesized according to statistics aggregated

[2]The following note is given at https://lehd.ces.census.gov/data/ for the data: "The data released by LEHD are based on tabulated and modeled administrative data, which are subject to error. Because the estimates are not derived from a probability-based sample, no sampling error measures are applicable. However, the data are subject to nonsampling errors, which can be attributed to many sources: misreported data, late reporters whose records are missing and imputed, and geographic/industry edits and imputations. The accuracy of the data is impacted by the joint effects of these nonsampling errors. While no direct measurement of these joint effects has been obtained, precautionary steps are taken in all phases of collection and processing to minimize the impact of nonsampling errors."

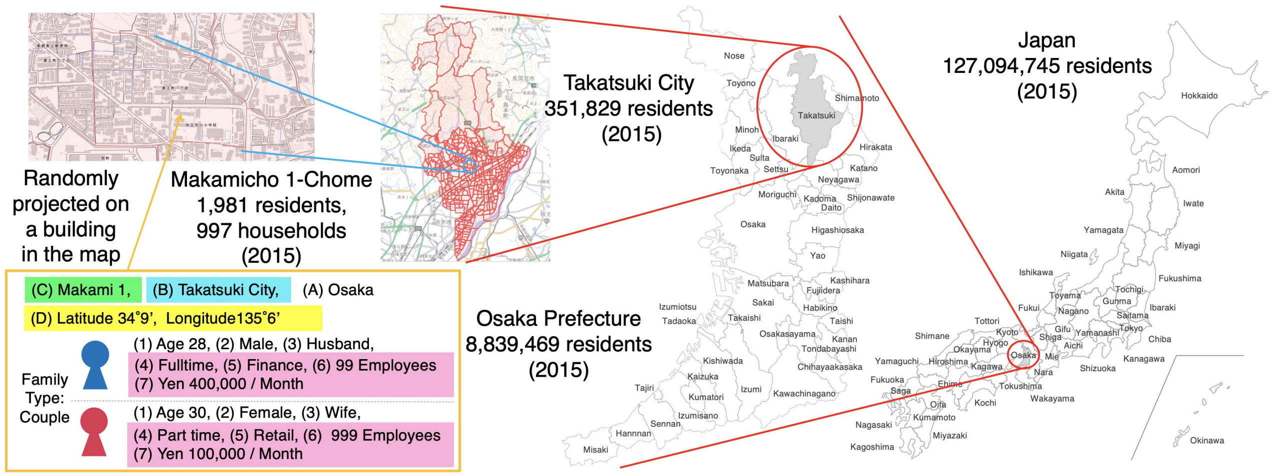


Fig. 2. Address and attributes of household members (small area map by geoshape local government ID dataset).

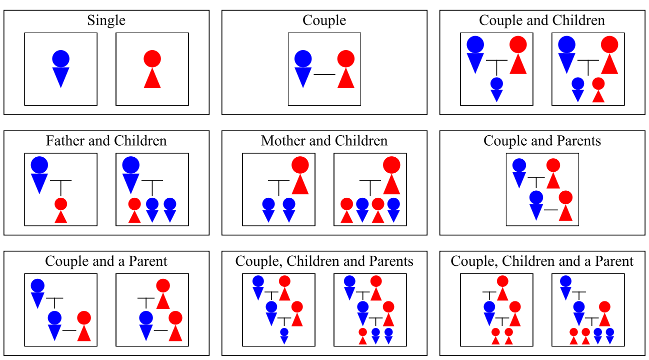


Fig. 3. Nine family types.

by prefecture and by city, town, or village [6]. To assign households to small areas in a city, we utilize the number of residents by sex in each small area [7]. Then, we allocate each household to a building in the assigned small area to specify the address attributes (C) and (D) in Fig. 2 [7]. Attributes (4)–(7) are assigned to each worker based on Basic Survey on Wage Structure [8].

B. Workplace Assignment From City to City

The workplace assignment is conducted to each worker of a household in a synthetic population. Using three attributes, such as (1) age, (2) sex, and (5) working industry, we assign a workplace to each worker.

To assign a workplace to each worker, we need ODI data in Fig. 1. In Japan, we have ODI data from residential cities to working cities in the census data. However, incomplete destination information is only available in some cities, towns, and villages with less than 200 000 residents for privacy reasons. To compensate for the lack of finer destinations, we propose a method for all cities, towns, and villages (especially for cities, towns, and villages with less than 200 000 residents).

1) All Cities, Towns, and Villages: Statistics Bureau of Japan releases ODI data for all cities, towns, and villages in a restricted manner. Table I shows the number of workers by destinations and by working industry in an originating city (residential city of workers). Table I indicates the destinations, not in a finer way, but in a vague way, such as "In the Same Prefecture" or "In Other Prefectures." Therefore, we employ Table II, which indicates the number of workers by destination at the city level. From Table II, we can see the number of male and female workers in each city, town, or village in Japan. It should be noted that Table II does not indicate the industry of workers. Among 1896 wards, cities, towns, and villages in Japan, 487 are listed in Table II for workers in Takatsuki, Osaka, Japan, in 2015 (with four more destinations, such as "At Home," "In City," "Unknown in Japan or Outside Japan," and "Not Available"; the number of destinations becomes 491 in total).

Before assigning cities of workplaces to workers, we adjust the number of workers in "T) Industry unable to classify" in Table I since the respondents classified in T) miswrote their occupations that are unable to classify. On the other hand, Economic Census that is used in Section III does not have such categories since all companies and self-employers register themselves as one of A)–S). Table I shows that there are 5622 male workers in T). The same statistics show that there are 4644 female workers in T). Therefore, there are 10 266 workers in T) in total. We distribute these workers to industries A)–S) according to the rate of male and female workers in each industry over A)–S).

Using Tables I and II, we assign a workplace using the following procedure.

Step 1: Assign one of six workplace categories a in Table I according to the sex x (male or female) and the industry d of the workers in a synthetic population [attributes (2) and (5) in Fig. 2]. Let us denote $n(x, d, a)$ as the number of x workers who work for industry d in one of the area category a in Table I. The workplace category is defined using the following rate:

$$n(x, d, a) \Big/ \sum_{a \in A} n(x, d, a) \tag{1}$$

where A is a set of six workplace categories in Table I, such as "At home," "In the same city," "In another city in the same prefecture," "In another city in another prefecture," "Unknown in Japan or outside Japan," and "Not available." Note that the

TABLE I

NUMBER OF MALE WORKERS BY DESTINATIONS AND BY INDUSTRY (AGE 15 OR MORE, TAKATSUKI, OSAKA IN 2015)

Industry Type	At Home	In City	In the Same Prefecture	In Other Prefectures	Unknown in Japan or Outside Japan	Not Available	Total
A) Agriculture & Forestry	286	199	52	29	3	2	571
B) Fisheries	-	-	-	2	-	-	2
C) Mining	-	13	1	6	-	-	20
D) Construction	923	2,443	3,094	974	178	131	7,743
E) Manufacturing	285	4,158	7,725	4,238	144	95	16,645
F) Electricity, Gas, Heat Supply & Water	-	122	339	129	1	1	592
G) Information & Communications	141	156	2,447	684	32	16	3,476
H) Transport & Postal Services	117	2,375	3,848	1,032	91	61	7,524
I) Wholesale & Retail Trade	633	3,181	3,848	1,950	93	98	11,872
J) Finance & Insurance	61	125	1,202	561	18	6	1,973
K) Real Estate, Good Rental & Leasing	373	663	1,063	294	15	26	2,434
L) Scientific, Professional, & Technical Services	545	471	1,651	708	33	14	3,422
M) Accommodations & Eating and Drinking Serv.	136	1,223	1,034	440	35	35	2,903
N) Living and Personal & Amusement Services	228	813	716	204	14	13	1,988
O) Education, Learning Support	73	1,026	1,529	670	5	8	3,311
P) Medical, Health Care & Welfare	205	2,428	1,849	541	13	24	5,060
Q) Compound Services	-	274	111	63	-	2	450
R) Other Services	326	2,059	2,659	746	59	50	5,899
S) Government	2	727	1,201	470	7	5	2,412
T) Industry Unable to Classify	242	750	870	289	197	3,274	5,622
Total	4,576	23,206	37,308	14,030	938	3,861	83,919

TABLE II

NUMBER OF MALE AND FEMALE WORKERS BY 491 DESTINATIONS (AGE 15 OR MORE, TAKATSUKI IN 2015)

Industry Type	Male	Female
At Home	4,576	3,787
In City	23,206	33,937
Chuo-ward, Sapporo-city, Hokkaido	2	1
...	...	...
Onna-son, Okinawa	-	1
In Japan or Outside Japan	938	322
Unknown	3,861	2,596
Total	83,919	66,295

workplace is not assigned stochastically but assigned randomly up to $n(x, d, a)$.

Step 2: Assign a city, town, or village in the worker's living prefecture according to Table II up to the number of workers in the specific industry. Let us denote $n(x, w)$ as the number of x workers who work in a city, town, or village w. One of cities, towns, and villages is assigned using the following rate:

$$n(x, w) / \sum_{w \in B} n(x, w) \tag{2}$$

where B is a set of workplaces in the living prefecture in Table II (e.g., 52 male workers in A) Agriculture & Forestry in Table I).

Step 3: Assign a city, town, or village in other prefecture up to the number of workers in the industry (e.g., 29 male workers in agriculture and forestry in Table I). One of cities, towns, and villages is assigned using the following rate:

$$n(x, w) / \sum_{w \in C} n(x, w) \tag{3}$$

where C is a set of workplaces in cities of another prefecture in Table II.

Step 4: The rest of the workers are not assigned to any city, town, or village since their workplace is "Unknown in Japan or Outside Japan" or "Not available" (e.g., three in "Unknown in Japan or Outside Japan" and two in "Not available" in Agriculture and Forestry in Table I).

2) Cities, Towns, and Villages More Than 200 000: Statistics Bureau of Japan releases more precise ODI data for cities with more than 200 000 residents. Those cities have the statistics that connect Tables I and II. Table III shows an example of Takatsuki city. Although Tables I and II distinguish the sex, Table III shows the total workers of male and female workers (i.e., no distinction in sex) in every industry type due to privacy reasons. The number of cities, towns, and villages is 487 in Table III for Takatsuki city, the same as Table II. In Japan, the number of wards and cities with more than 200 000 is 290 wards and cities among 1896. The total number of workers in these 290 bodies becomes 52.2% of all workers in Japan.

Using Table III without the distinction of sex, we assign a workplace using the following procedure for cities with more than 200 000 residents.

Step 1: Assign one of 491 destinations in Table III according to the industry of the workers in a synthetic population [attributes (2) and (5) in Fig. 2]. Let us denote $n(x, d, w)$ as the number of x workers who work in a city, town, or village w. One of cities, towns, and villages is assigned using the following rate:

$$n(x, d, w) / \sum_{a \in D} n(x, d, w) \tag{4}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE III

NUMBER OF WORKERS BY DESTINATIONS AND BY INDUSTRY (AGE 15 OR MORE, TAKATSUKI, OSAKA IN 2015)

Industry Type	At Home	In City	Chuo-ward, Sapporo, Hokkaido	...	Onna-Son Okinawa	Unknown in Japan or Outside Japan	Not Available	Total
A) Agriculture & Forestry	423	246	-	...	-	4	2	778
B) Fisheries	-	-	-	...	-	-	-	2
C) Mining	1	13	-	...	-	-	-	21
D) Construction	1,233	2,945	1	...	-	184	144	9,120
E) Manufacturing	540	7,441	1	...	-	165	137	23,263
F) Electricity, Gas, Heat Supply & Water	1	179	-	...	-	1	2	698
G) Information & Communications	241	246	-	...	-	36	21	4,475
H) Transport & Postal Services	161	3,437	-	...	-	98	72	9,499
I) Wholesale & Retail Trade	1,165	10,157	1	...	-	158	187	24,542
J) Finance & Insurance	97	628	-	...	-	30	15	4,340
K) Real Estate, Good Rental & Leasing	711	1,166	-	...	-	24	33	3,789
L) Scientific, Professional, & Technical Services	814	822	-	...	-	39	22	5,113
M) Accommodations & Eating and Drinking Serv.	340	4,143	-	...	-	49	80	7,321
N) Living and Personal & Amusement Services	560	2,117	-	...	1	24	28	4,731
O) Education, Learning Support	372	3,104	-	...	-	27	24	7,913
P) Medical, Health Care & Welfare	464	13,001	-	...	-	40	82	20,330
Q) Compound Services	1	497	-	...	-	-	2	757
R) Other Services	552	3,992	-	...	-	83	81	9,600
S) Government	2	1,356	-	...	-	8	6	3,674
T) Industry Unable to Classify	685	1,653	-	...	-	290	5519	10,266
Total	8,363	57,143	3	...	1	1,260	6,457	150,214

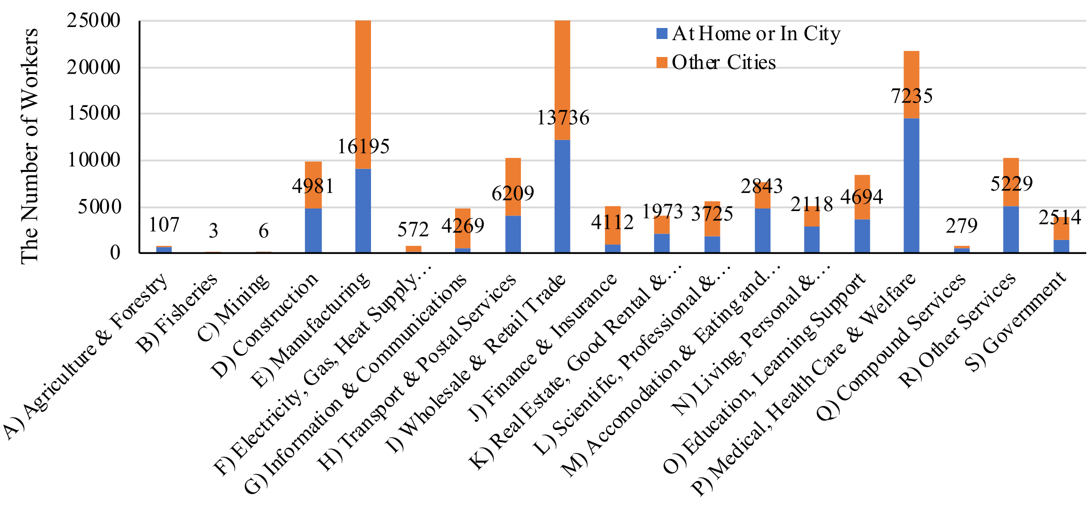


Fig. 4. Number of workers by industry in Takatsuki, Osaka, Japan, in 2015.

where D is a set of workplaces in cities of another prefecture in Table III.

Step 2: The rest of the workers are not assigned to any city since their workplace is "Unknown in Japan or Outside Japan" or "Not available" (e.g., four in "Unknown in Japan or Outside Japan" and two in "Not available" in Agriculture & Forestry in Table III).

3) Differences in City Allocation Between Two Methods: Since Tables I and II are available in every city in Japan, we apply the proposed workplace assignment method in Section II-B1 to Takatsuki city, Osaka, Japan. Fig. 4 shows the number of workers by industry type based on Table I (i.e., Step 1 in Section II-B1) after distributing the number of workers in T) to each industry for Takatsuki city. In each industry, a numeral figure over a bar indicates the number of workers in other cities, towns, and villages. To each of those workers, one of the cities, towns, and villages in Table II is assigned.

Fig. 5 shows the number of workers assigned to cities correctly and mistakenly. We apply the proposed method using Tables I and II according to the procedure in Section II-B1. We also applied another method using Table III by the procedure in Section II-B2. We assigned cities to workers using ten sets of the synthesized populations of Takatsuki city. In Fig. 5, the solid bar shows the average number of workers assigned to the correct cities indicated in Table III. The slashed bar shows the number of workers assigned mistakenly to cities by the proposed method. Figures on the top of the bars indicate

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

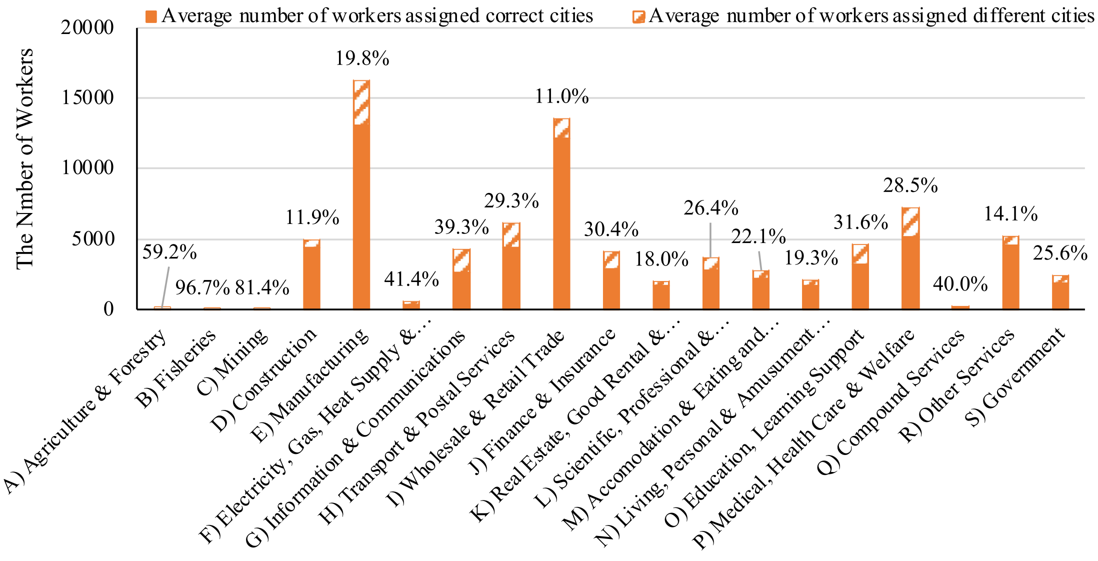


Fig. 5. Number of workers correctly and differently assigned to cities to work.

TABLE IV
ACCURACY IN EACH INDUSTRY BY THE PROPOSED METHOD

Industry Type	Accuracy	Rate of Home City
A) Agriculture & Forestry	92.2%	86.8%
B) Fisheries	3.3%	0%
C) Mining	74.1%	68.2%
D) Construction	94.0%	49.4%
E) Manufacturing	87.3%	35.6%
F) Electricity, Gas, Heat Supply & Water	69.4%	26.2%
G) Information & Communications	65.5%	12.2%
H) Transport & Postal Services	82.3%	39.7%
I) Wholesale & Retail Trade	94.2%	47.5%
J) Finance & Insurance	75.0%	17.7%
K) Real Estate, Good Rental & Leasing	91.2%	51.1%
L) Scientific, Professional, & Technical Services	82.3%	33.2%
M) Accommodations & Eating and Drinking Serv.	91.8%	63.0%
N) Living and Personal & Amusement Services	91.9%	57.9%
O) Education, Learning Support	82.5%	44.6%
P) Medical, Health Care & Welfare	90.5%	66.8%
Q) Compound Services	86.4%	66.0%
R) Other Services	92.8%	49.0%
S) Government	84.0%	37.3%
Total	88.2%	46.3%

the rate of workers mistakenly assigned among all workers who work in other cities in each industry.

Table IV shows the classification rate of workers who are correctly assigned to the cities of workplaces by the proposed method and the rate of workers who are working in their home city. The accuracy of workplace assignment by the proposed method is 88.2% for all industries. The figures with underline in the tables show that the value is larger than the overall value (i.e., 88.2% in "Accuracy" and 46.3% in "Rate of Home City"). This result shows that the industry with higher accuracy than the overall accuracy comes from the higher rate of workers working in their home city.

Fig. 6 depicts the relations between the number of all workers by industries and the number of correctly assigned workers by the proposed method. The vertical difference between each maker and the diagonal line shows the number of workers who are mistakenly assigned to workplaces. From Fig. 6, the proposed method using only the OD data assigns relatively many workers correctly to their workplaces. Of course, this depends on the number of workers who are working outside their living city. For Takatsuki-city, the workers are working in their living city are 43.6% whereas that in the entire Japan are 53.8%. That is, the accuracy of workplace assignment becomes higher than the result of Takatsuki-city for the entire Japan.

III. Workplace Assignment at Small Area Level

Using the procedures in Section II, we have assigned a working city to each worker. In this section, we assign a small area to each worker in the assigned city. As shown in Fig. 2, each city, town, or village has small areas. Census data are accumulated based on these small areas. We employ Economic Census for Business Frame in Japan. This census is a survey targeting all establishments and enterprises to identify the structures and activities of businesses in Japan and is conducted every five years. Although we employ census data in 2015 to synthesize a population and assign cities of workplaces, Economic Census was not conducted in 2015. We employ Economic Census in 2014 in this article.

A. Location of Workplace Who Work at Home

We assign the home coordinate (i.e., latitude and longitude) as the workplace coordinate for workers who work at home in Tables I or III. We do not need the results of Economic Census for them.

B. Location of Workplace Who Work in a Specific City

For each worker who works within the city where it lives or in other city assigned by the procedures in Section II, we employ Economic Census. In each city, town, or village in Japan, Economic Census shows the number of workers in each

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE V
NUMBER OF WORKERS BY SMALL AREA AND BY INDUSTRY (AGE 15 OR MORE, IBARAKI, OSAKA IN 2014)

Industry Type	Misaki-cho	Mishimaoka 1-chome	Mishimaoka 2-chome	...	Nonomiya 2-chome	Total
A) Agriculture & Forestry	-	-	-	...	4	233
B) Fisheries	-	-	-	...	-	-
C) Mining	-	-	-	...	-	6
D) Construction	18	11	6	...	56	5,113
E) Manufacturing	21	181	2	...	97	11,562
F) Electricity, Gas, Heat Supply & Water	-	-	-	...	15	308
G) Information & Communications	-	-	5	...	-	350
H) Transport & Postal Services	138	13	3	...	127	10,543
I) Wholesale & Retail Trade	19	97	18	...	67	24,506
J) Finance & Insurance	-	18	-	...	4	3,415
K) Real Estate, Good Rental & Leasing	4	-	8	...	25	4,117
L) Scientific, Professional, & Technical Services	-	-	2	...	1	3,728
M) Accommodations & Eating and Drinking Serv.	13	20	2	...	10	9,007
N) Living and Personal & Amusement Services	5	1	5	...	-	3,880
O) Education, Learning Support	-	-	1	...	-	7,459
P) Medical, Health Care & Welfare	18	-	50	...	-	17,828
Q) Compound Services	-	-	5	...	-	324
R) Other Services	26	3	-	...	25	6,893
S) Government	-	-	-	...	-	2,224
Total	262	344	107	...	431	111,496

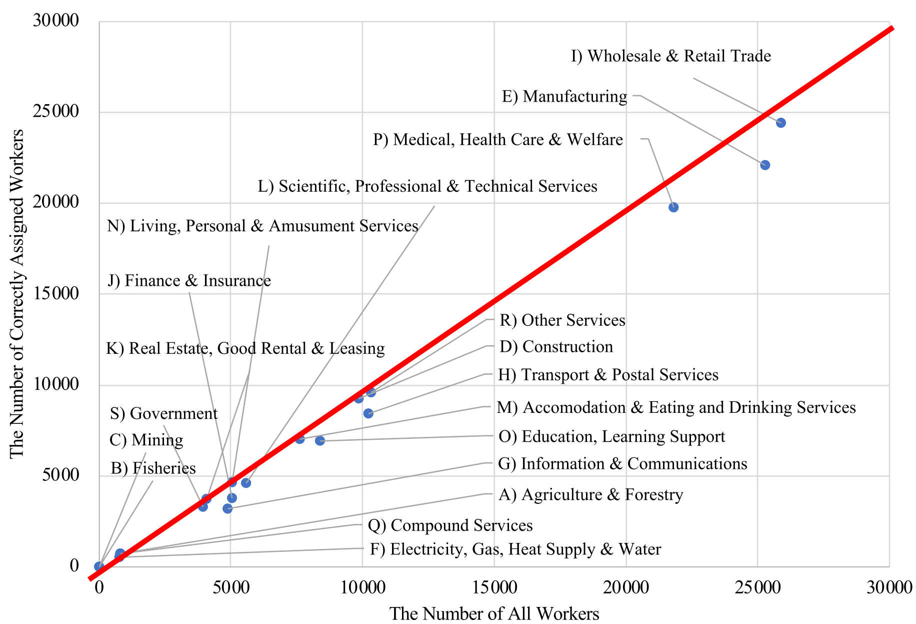


Fig. 6. Number of all workers by industries and the number of correctly assigned workers.

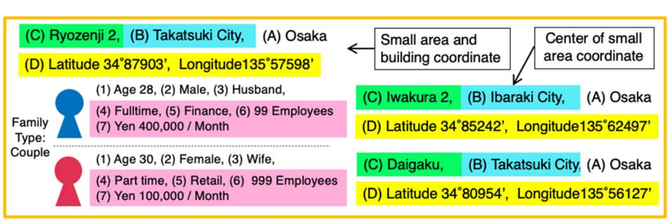


Fig. 7. Attributes of household members and workplace coordinate of workers in the household (pseudo household is shown).

small area by industry. Table V shows an example of Ibaraki city, the western neighbor to Takatsuki, Osaka, Japan. One of the small areas is assigned to a worker of the synthetic population according to the following probability based on Table V:

$$f_{w,d}(c_k) = n(w,d,c_k) \Big/ \sum_{i=1}^{N_w} n(w,d,c_i), \quad (k = 1, 2, \ldots, N_w) \tag{5}$$

where $n(w, d, c_k)$ is the number of workers who work for a company of industry d in a small area c_k, a city, town, or village w, and N_w is the number of small areas in w. It should be noted that the number of assigned workers to each small area does not always become equal to the corresponding value in the table, such as Table IV, since a small area is assigned

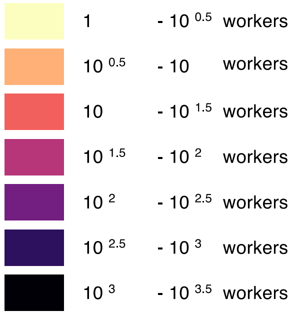


Fig. 8. Color scales of the number of workers in each small area from Takatsuki, Osaka, Japan.

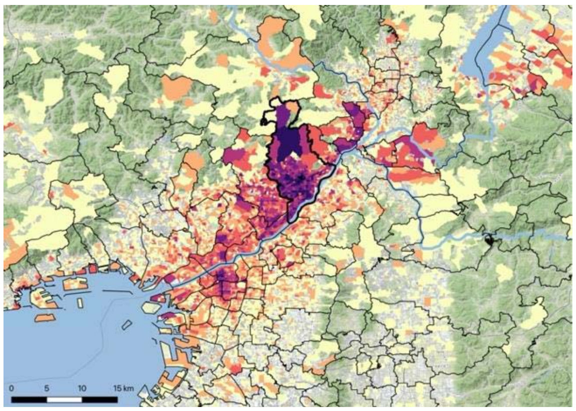


Fig. 9. Number of workers in all industries (A–S).

probabilistically to each worker. Furthermore, workers from city, town, or village with less than 200 000 residents are sometimes assigned to a workplace city mistakenly, as shown in Fig. 5. Therefore, there is room to optimize the number of workers after assigning workers from all originating cities, towns, and villages. This challenge remains for further study.

After assigning one of the small areas in the assigned city (town or village) to each worker, we assign a center coordinate (i.e., latitude and longitude) of the small area in Economic Census. Using this coordinate developers of agent-based simulations easily deploys movements of each agent from home to workplace. It is not a coordinate of a building of a workplace, although a coordinate for each home [(D) in Fig. 2] is represented by each building coordinate. Fig. 7 shows the workplace attributes for a household with workers. The coordinate of the center of the small area where the worker's workplace exists is added to each worker. The coordinate of the center of the small area is given up to the fifth decimal. On 35° north latitude, the difference of 0.00001° in longitude becomes 0.913 m. It is sufficiently precise since each building for residents or workers is larger than 1 m^2.

IV. Results of Workplace Assignment

We depict the number of workers in some industries on the map of Osaka Prefecture and its neighboring areas (Hyogo, Kyoto, Nara, and Shiga Prefectures). Fig. 8 shows the color scales of the number of workers from Takatsuki city, Osaka, Japan. Fig. 9 shows the number of workers in all industries

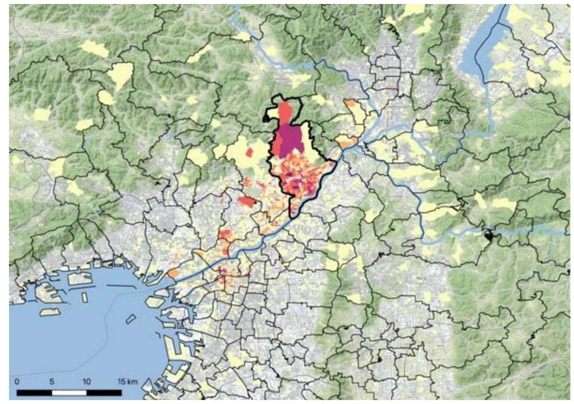


Fig. 10. Number of workers in D) Construction.

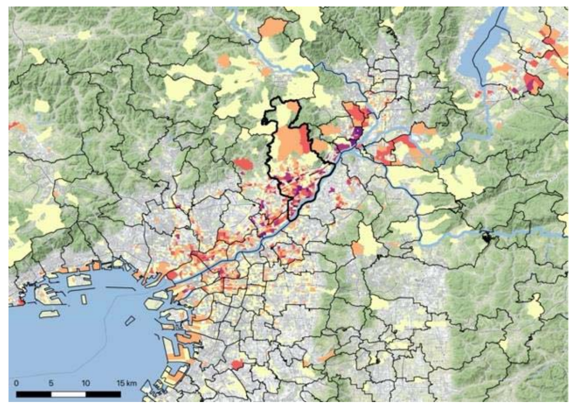


Fig. 11. Number of workers in E) Manufacturing.

from Takatsuki city in each small area. We show the number of workers in D) Construction, E) Manufacturing, H) Transport and Postal Services, and P) Medical and Welfare Services. From Figs. 8 and 9, we can see that those residents in Takatsuki-city commute to Hyogo Prefecture (Western side of Osaka), Kyoto Prefecture (Northeast from Takatsuki-city), and Shiga Prefecture (Eastern side of Kyoto, including Biwa Lake). On the map, the Yodo River flows from Biwa Lake to Osaka Bay from Northeast to Southwest. Since Takatsuki-city locates on the northern side of the river and two train lines (Japan Railways and Hankyu Railways) run along with the river, we can see more workers on the northern side of the river than on the southern side of Takatsuki.

From Figs. 10 and 11, the workers in D) Construction and E) Manufacturing are working along with the river and downtown Osaka (the plain next to Osaka Bay). Those workers commute to the southern side of the river in downtown. From Fig. 12, the workers in H) Transport and Postal Services are working in areas along the northern side of the river. Since many storehouses are located along the river, many workers in transport and postal services are working in these areas. Fig. 13 shows a different tendency of workplaces in P) Medical and Welfare Services. Workers in this category reside in the same city. Since they may have a night shift for their work, they live and work in the same city.

In order to see the fact of tendency observed in Fig. 13, we accumulate the ratio of workers living and working in the

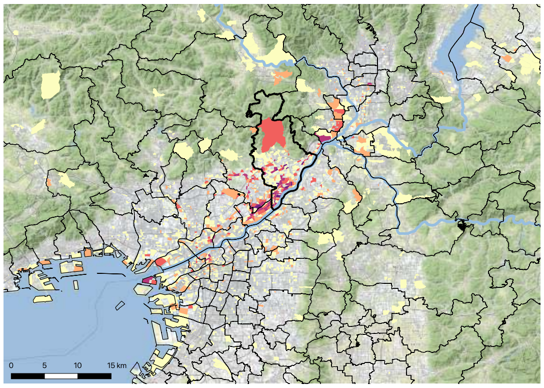


Fig. 12. Number of workers in H) Transport & Postal Services.

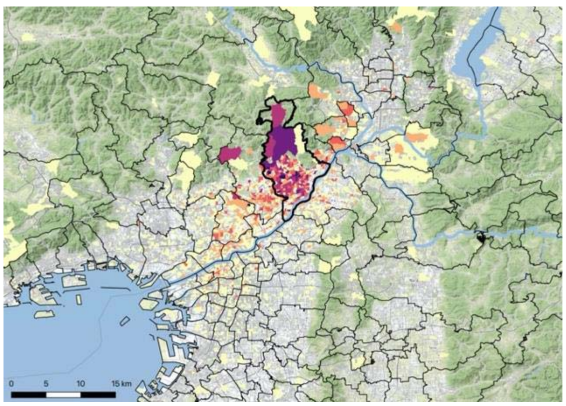


Fig. 13. Number of workers in P) Medical, Health Care & Welfare.

Fig. 14. Four cities on the north side of the Yodo river.

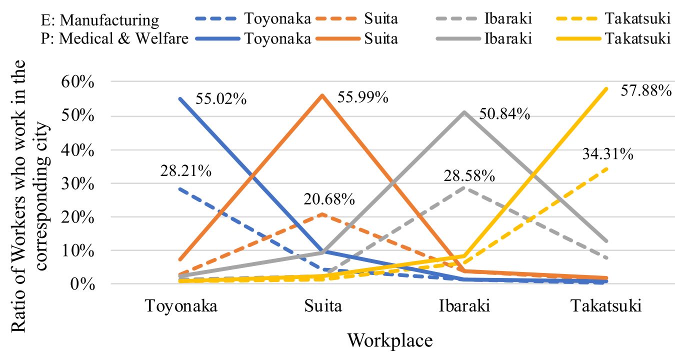


Fig. 15. Ratio of workers who live and work in the same city.

same city in four cities, such as Takatsuki, Ibaraki, Suita, and Toyonaka, in Figs. 14. Fig. 15 shows the ratio of workers living and working in the same city in E) Manufacturing and P) Medical, Health Care & Welfare services. Dashed and solid lines indicate the ratio of workers who live in the corresponding city. The horizontal axis shows the city of the workplaces. From Fig. 15, we can see the ratio of workers living and working in the same city. The ratio of workers in E) Manufacturing work in the same city from their living city up to 34.31%. On the other hand, the ratio of workers in P) Medical and Welfare Services becomes over 50% among these four cities. We can see the apparent tendency of workers of P) Medical and Welfare Services to work in their home city.

V. Conclusion

This article proposed a method to assign workplace attributes to each worker of the synthetic population in all cities, town, and villages in Japan. Since the ODI data for cities, towns, and villages with less than 200 000 residents have only restricted information on the destination city for work, we proposed a method to assign a workplace using OD data without the industry information. By applying the proposed method to Takatsuki-city, Osaka, Japan, we showed the difference between the number of assigned workers using the complete ODI data and the number of assigned workers by the proposed method with the restricted ODI data. We show that 88.2% of workers in a city in Japan are assigned to correct cities as workplaces by our proposed method. Since Takatsuki-city has a larger number of workers who are working outside the city than the average of the entire Japan, we can expect the better workplace assignment in many places in Japan.

We have already developed our proposed method to apply to all cities in Japan. We will soon release the data to research communities to develop social simulations. Synthetic population data with the workplace can be utilized in many applications, such as epidemic simulations and transportation design in target regions or areas. It is also needed to estimate the number of victims of earthquakes or floods and plan to rescue them from places struck by disasters.

References

[1] M. S. Narassima et al., "An agent-based model to assess coronavirus disease 19 spread and health systems burden," *Int. J. Elect. Comput. Eng.*, vol. 12, no. 4, pp. 4118–4128, Aug. 2022, doi: 10.11591/ijece.v12i4.pp4118-4128.

[2] M. Renardy, M. Eisenberg, and D. Kirschner, "Predicting the second wave of COVID-19 in Washtenaw County, MI," *J. Theor. Biol.*, vol. 507, Dec. 2020, Art. no. 110461, doi: 10.1016/j.jtbi.2020.110461.

[3] K. El Emam, L. Mosquera, E. Jonker, and H. Sood, "Evaluating the utility of synthetic COVID-19 case data," *JAMIA Open*, vol. 4, no. 1, pp. 1–10, Mar. 2021.

[4] N. Popper et al., "Synthetic reproduction and augmentation of COVID-19 case reporting data by agent-based simulation," *Data Sci. J.*, vol. 20, no. 16, pp. 1–13, Apr. 2021, doi: 10.1101/2020.11.07.20227462.

[5] F. Wood et al., "Planning as inference in epidemiological dynamics models," *Frontiers Artif. Intell.*, vol. 4, Mar. 2022, Art. no. 550603, doi: 10.3389/frai.2021.550603.

[6] H. Nagai and S. Kurahashi, "Measures to prevent and control the spread of novel coronavirus disease (COVID-19) infection in tourism locations," *SICE J. Control, Meas., Syst. Integr.*, vol. 15, no. 2, pp. 1–12, 2022, doi: 10.1080/18824889.2021.2012398.

[7] W. D. Wheaton et al., "Synthesized population databases: A US geographical database for agent-based models," *Methods Rep.*, vol. 2009, no. 10, p. 905, 2009.

[8] A. Adiga et al., "Generating a synthetic population of the United States," Network Dyn. Simul. Sci. Lab., Virginia Bioinf. Inst., Virginia, Tech. Rep. NDSSL 15-009, pp. 1–9, Jan. 2015.

[9] J. Robards, C. Gale, and D. Martin, "Creating a synthetic spatial microdata set for zone design experiments using 2011 Census and linked administrative data," *Proc. Geograph. Inf. Sci. Res. UK*, vol. 2017, no. 98, pp. 1–6, 2017.

[10] T. Murata, T. Harada, and D. Masui, "Comparing transition procedures in modified simulated-annealing-based synthetic reconstruction method without samples," *SICE J. Control, Meas., Syst. Integr.*, vol. 10, no. 6, pp. 513–519, Nov. 2017, doi: 10.9746/jcmsi.10.513.

[11] T. Harada and T. Murata, "Projecting households of synthetic population on buildings using fundamental geospatial data," *SICE J. Control, Meas., Syst. Integr.*, vol. 10, no. 6, pp. 505–512, Nov. 2017, doi: 10.9746/jcmsi.10.505.

[12] T. Murata, S. Sugiura, and T. Harada, "Income allocation to each worker in synthetic populations using basic survey on wage structure," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Nov. 2017, pp. 471–476, doi: 10.1109/SSCI.2017.8285242.

[13] L. Willem and P. J. K. Libin. (Feb. 1, 2021). *Synthetic Population Data for Belgium for STRIDE*. Zenodo. YdmUZS33KCM, [Online]. Available: https://zenodo.org/record/4485995#, doi: 10.5281/ZENODO.4485995.

[14] A. G. Wilson and C. E. Pownall, "A new representation of the urban system for modeling and for the study of micro-level interdependence," *Area*, vol. 8, no. 4, pp. 246–254, 1976.

[15] W. E. Deming and F. F. Stephan, "A least squares adjustment of a sampled frequency table when the expected marginal totals are known," *Ann. Math. Statist.*, vol. 11, pp. 428–444, Dec. 1940.

[16] J. Barthelemy and P. L. Toint, "Synthetic population generation without a sample," *Transp. Sci.*, vol. 47, no. 2, pp. 266–279, May 2013.

[17] F. Gargiulo, S. Ternes, S. Huet, and G. Deffuant, "An iterative approach for generating statistically realistic populations of households," *PLoS ONE*, vol. 5, no. 1, p. e8828, 2010, doi: 10.1371/journal.pone.0008828.

[18] M. Lenormand and G. Deffuant, "Generating a synthetic population of individuals in households: Sample-free vs sample-based methods," *J. Artif. Societies Social Simul.*, vol. 16, no. 4, pp. 1–9, 2013.

[19] M. M. M. Abdel-Aal, "Calibrating a trip distribution gravity model stratified by the trip purposes for the city of Alexandria," *Alexandria Eng. J.*, vol. 53, pp. 677–689, Sep. 2014.

[20] N. Fournier, E. Christofa, A. P. Akkinepally, and C. L. Azevedo, "Integrated population synthesis and workplace assignment using an efficient optimization-based person-household matching method," *Transportation*, vol. 48, no. 2, pp. 1061–1087, Apr. 2021, doi: 10.1007/s11116-020-10090-3.

[21] P. Ye, F. Zhu, S. Sabri, and F.-Y. Wang, "Consistent population synthesis with multi-social relationships based on tensor decomposition," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 5, pp. 2180–2189, May 2020.

Tadahiko Murata (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees from Osaka Prefecture University, Osaka, Japan, in 1994, 1996, and 1997, respectively.

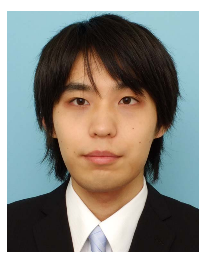
He was the Director of the Policy Grid Computing Laboratory, Kansai University, Osaka, from 2005 to 2010. He was a Visiting Scholar with the University of Chicago, Chicago, IL, USA, from 2010 to 2011. He is currently a Professor with Kansai University. His current research interests include social simulation, parallel and distributed computing, soft computing, and multiobjective optimization.

Dr. Murata received the Best Presentation Award from the Institute of Systems, Control and Information Engineers (ISCIE) in 1997; the Best Paper awards from Special Interest Group (SIG) on Social Systems, Society for Instruments and Control Engineers (SICE), in 2017; Meritorious Service Award from IEEE SMCS in 2019; the Best Paper Award from SICE in 2019; and the Best Demonstration Award from SOFT in 2020. He belongs to the Japanese Society for Evolutionary Computation (JPNSEC), the Japanese Society for Artificial Intelligence (JSAI), SICE, the ISCIE, and the Japan Society for Fuzzy Theory and Intelligent Informatics (SOFT). He is currently the Vice President of the Organization and Planning, IEEE SMCS. He was an Associate Editor of the IEEE TRANSACTION ON SYSTEMS, MAN, AND CYBERNETICS—Part B and is a Senior Editor of *Review of Socionetwork Strategies* (Springer, Tokyo), the President of JPNSEC, and the Editor-in-Chief of *Transactions of JPNSEC*.

Daiki Iwase (Student Member, IEEE) received the B.S. degree from Kansai University, Osaka, Japan, in 2022.

He is currently an Information Systems Architect with JR East, Tokyo, Japan.

Mr. Iwase received the Best Presentation Award from Special Interest Group (SIG) on Social Systems, Society for Instruments and Control Engineers (SICE), in 2022. He was a Student Member of the SICE from 2021 to 2022.



Takuya Harada (Member, IEEE) received the B.S., M.S., and Ph.D. degrees from Kansai University, Osaka, Japan, in 2015, 2017, and 2018, respectively.

He was with Aoyama Gakuin University, Tokyo, Japan, from 2019 to 2022 and joined the Shibaura Institute of Technology, Tokyo, as an Assistant Professor, in 2022. His current research interests include social simulation, parallel and distributed computing, and image recognition.

Dr. Harada received the Best Paper Award from Special Interest Group (SIG) on Social Systems, Society for Instruments and Control Engineers (SICE), in 2017, the Best Paper Award from SICE in 2019, and the Best Presentation Award from Systems and Information Department, SICE, in 2020. He belongs to the SICE.