Workplace Assignment Optimization Method Considering Multiple Municipalities Using Synthetic Population*

Shunta Horigami, and Tadahiko Murata, *Fellow, IEEE*

***Abstract*—In this paper, we improve a workplace assignment method considering multiple inflow municipalities. The former method assigns workplaces in multiple municipalities to workers living in one municipality. Although workers in the target municipality are assigned to multiple municipalities, small areas within the municipalities are stochastically assigned to workers. There may happen the difference between working distribution by stochastic assignment and real working distribution in the municipalities of workplaces. In this paper, we focus a municipality and assign workplaces in the target municipality to workers living in other municipalities. After assignment, we find the difference between the number of workers in each small area in the target municipality and the real distribution of workers in the target municipality. We propose a method to adjust the number of workers in each small area in the target municipality. Our simulation result clearly shows that the difference of the number of workers is reduced by the proposed method.**

I. Introduction

In this paper, we improve a workplace assignment method by Murata et al. [1] considering multiple inflow municipalities. The former method [1] assigns workplaces in multiple municipalities to workers living in one municipality. Although workers in the target municipality are assigned to multiple municipalities, small areas within the municipalities are stochastically assigned to workers. There may happen the difference between working distribution by stochastic assignment and real working distribution in the municipalities of workplaces.

Origin-destination (OD) surveys are employed for workplace estimation of workers by many researchers. Those estimation methods can be categorized in two ways: One employs (1) sampling surveys and (2) complete surveys. (1) sampling surveys extract some subjects of all population. On the other hand (2) complete surveys target all workers in the population.

As an example of employing (1) sampling OD surveys, Abdel-Aal [2] assign workplaces by 15 zones to workers in Alexandria, Egypt using a sampling OD survey. They employed the IPF (Iterative Proportional Fitting) procedure to estimate the entire movements between 15 zones in Alexandria.

As examples of employing (2) complete OD surveys, Fournier et al. [3] utilized Longitudinal Employer Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES) collected by the Center for Economic Studies of the United States Census Bureau. Since there are no ODI (Origin-Destination-Industry) data in LODES, they combined workplace origin-destination totals (OD), workplace origin totals by industries (OI), and workplace destinations by industries (DI). As for LODES, the note is given such as "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." While complete data has advantage compared to probability-based sample, it is also noted that "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." We need to accept errors from above reasons.

Ye et al. [4] employed Chinese National Population and Economic Census. Chinese National Population Census is conducted in two ways: One targets whole population called Short Table, and the other target about 9.5% of the whole population for a detail survey called Long Table. They also employ the national economic investigation that targets whole target population in China. They utilize these complete survey and sampling survey such as Long Table.

Agriesti et al. [5] developed a workplace assignment method when only the number of workers in the districts is available. They applied their model in Estonia. Many related literature assignment of workplace are surveyed in this paper.

In Japan, there are complete ODI data for local municipalities with more than 200,000 residents. In those local municipalities, the number of residents whose origins (residential areas by small areas), destinations (workplaces by municipalities), and industries (type of working industries). As for local municipalities with less than 200,000 residents, the destination category is not fine as those with more than 200,000 residents. In the municipalities with less than 200,000 residents, 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," "In Japan or Outside Japan," and "Unknown." In Murata et al. [1],

* The Part of this research was was supported by "JST Mirai Program" (JPMJMI 23B1), "Joint Usage/Research Center for Interdisciplinary Large-scale Information Infrastructures (JHPCN)" in Japan (jh250064), HPCI Shared Storage (hp220354, hp230467, hp240468), and JSPS KAKENHI (25K05038).

Shunta Horigami is with the Graduate School of Information Science and Technology, the University of Osaka, Suita, Osaka 565-0871 Japan. (e-mail: horigami.syunta@ais.cmc.osaka-u.ac.jp)

Tadahiko Murata is with D3 Center, the University of Osaka, Ibaraki Osaka 567-0047 Japan. (phone: +81-06-6879-8790; e-mail: tadahiko.murata@osaka-u.ac.jp).

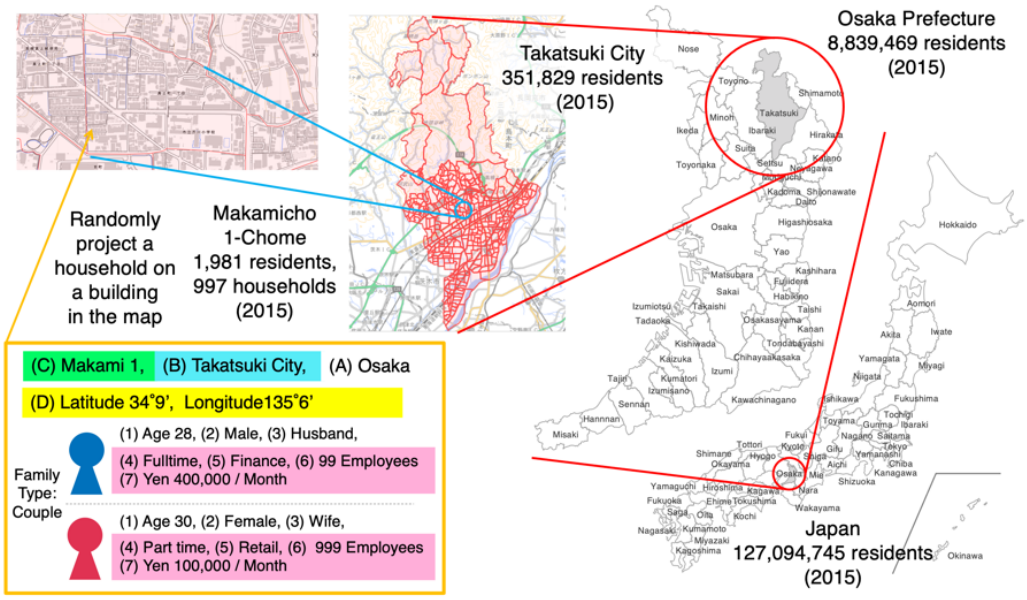

Figure 1. Address and attributes of household members (Small Area Map by Geoshape Local Government ID Dataset).

they proposed a workplace assignment method for workers in all cities, towns, and villages with populations less than 200,000.

Although appropriate assignment was done by their method, the assignment in each small areas in each municipality of workplace should be improved since workplace assignment within the municipality is stochastically done. That is, the number of workers in each small area may become different to the number of workers in the survey results.

In this paper, we focus a municipality and assign workplaces in the target municipality to workers living in other municipalities. After assignment, we find the difference between the number of workers in each small area in the target municipality and the real distribution of workers in the target municipality. Then, we propose a method to adjust the number of workers in each small area in the target municipality. Our simulation result clearly shows that the difference of the number of workers is reduced by the proposed method.

II. Synthetic Population

Fig. 1 (the left-bottom part) shows attributes for each household member in a synthetic population synthesized by Murata et al. [6, 7, 8]. There are four home address attributes of each household such as (A) prefecture, (B) municipality such as city, town or village, and (C) small area in the municipality, (D) latitude and longitude of the home. The National Census shows some statistics of households and individuals in (C) small areas. The census summarizes the statistics in the larger area such as (A) prefectures and (B) cities, towns or villages. As shown in **Fig. 1**, Japan has approximately 127 million residents in 2015. As a larger administration, there are 47 prefectures. As an example of a prefecture in Fig. 1, there are 33 cities, 9 towns, and 1 village with more than 8.8 million residents in Osaka Prefecture. Among those municipalities, only 8 cities have more than 200,000 residents. As an example of a municipality, Fig. 1 shows Takatsuki-city in Osaka. There are 448 small areas in Takatsuki-city and they are indicated by red lines in the figure. For each area, the fundamental geospatial data, including the shape data of each building with latitude and longitude released by Geospatial Information Authority of Japan, Ministry of Land, Infrastructure, Transport and Tourism, Japan. Murata et al. [7] projected each synthesized household on buildings in the map of the fundamental geospatial data [7].

Each household member in the synthesized household has seven biological and social attributes such as (1) age in one-year-age category, (2) sex such as male or female, (3) role in the household such as husband, wife, child or grandparent, (4) working status of fulltime or parttime, (5) working industry in 20 categories (19 industries and unknown), (6) industry size, and (7) income if it is working. The (3) role is assigned to each household member according to its family type, such as husband, wife, child, and parent. The (7) income is assigned by the method in [8]. One of nine family types shown in **Fig. 2** are also given to each household. 95% of all households in Japan are one of these nine family types.

III. Statistics Employed for Workplace Assignment

In Murata et al. [1], the workplaces are assigned using the statistics in **TABLE I**. Mainly they employed the statistics of National Census and Economic Census. In order to examine the number of workers in each municipality, Statistics E in Table 2 is a counter part of Statistics A in **TABLE I**. Statistics B' is a counterpart of Statistics B in the same table.

We employ Statistics E in **TABLE II** to find the name of municipalities where workers in the target municipality are living. In this paper, we take an example of Takatsuki-city in Osaka Prefecture that is depicted in **Fig. 1**. According to Statistics E, we found that over 93% of 109,557 workers who

work in Takatsuki-city are living in 67 cities, towns, and villages including Takatsuki-city itself.

The number of workers in Statistics D and E are different because they are conducted in different year. **Fig. 3** shows the difference of the number of workers by industry type in these two statistics. The alphabets of industry types are shown in **TABLE III**. Because there are some difference between two censuses, we adjust the number of workers in Statistics D using the following equation:

$$S_{s,i} = EC_{s,i} \cdot \frac{NC_i}{\sum_{i \in \{A,B,\ldots,S\}} NC_i}, \quad (1)$$

where $S_{s,i}$ is the adjusted number of workers of Industry i in Small Area s, $EC_{s,i}$ is the number of workers of Industry i in Small Area s in Economic Census, and NC_i is the number of workers of Industry i in the municipality in National Census.

We assign workplaces to workers who live in 67 municipalities by using the algorithm in [1] and obtain the distribution of workers in Takatsuki-city. We calculate the difference between the number of workers assigned by the former algorithm [1] and the number of workers adjusted by Equation (1) by each small area. The difference of assigned workers $Diff_{s,i}$ can be calculated using the following equation:

$$Diff_{s,i} = E_{s,i} - S_{s,i}, \quad (2)$$

where $E_{s,i}$ is the number of workers assigned by the former method [1] for Small area s and Industry i, and $S_{s,i}$ is the adjusted number of workers in Equation (1).

Fig. 4 shows a histogram of the number of combinations of Small area s and Industry i with $Diff_{s,i}$. **TABLE IV** shows the number of combinations of Small area s and Industry i in the absolute difference. From these figure and table, we can see that the difference becomes smaller than 10 in many combinations of Small area s and Industry i. By calculating (the total number of difference) /(the number of combinations of Small area s and Industry i), we can find that the difference of 4.22 workers averagely in each Small area s and Industry i.

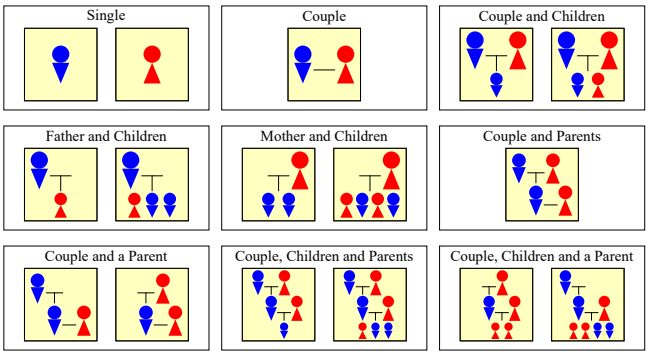

Figure 2. Nine family types.

TABLE I. THE ATTRIBUTES TO BE ASSIGNED AND EMPLOYED STATISTICS.

ID	Attributes to be assigned	Name of statistics	Municipalities included in statistics	Name of table
A	Workplace (Municipality)	National Census in 2017	21 large cities, Cities with prefectural government office	The number of workers by municipality of living place, by municipality of workplace, and by industry
B			All municipalities	The number of workers by municipality of living place, by municipality of workplace (nine categories), and by industry
C			All municipalities	The number of workers by municipality of workplace, and by living municipality
D	Workplace (Small area)	Economic Census in 2016	All municipalities	The number of offices and workers by sex in small areas in the municipality
-	Workplace (Coordinate)			Boundary data in Economic Census

TABLE II. THE STATISTICS TO EXAMINE THE ASSIGNMENT OF WORKPLACE.

ID	Attributes to be assigned	Name of statistics	Municipalities included in statistics	Name of table
E	Workplace (Municipality)	National Census in 2017	21 large cities, Cities with prefectural government office	The number of workers by municipality of workplace, by municipality of living place, and by industry
B'			All municipalities	The number of workers by municipality of living place, by municipality of workplace (nine categories), and by industry

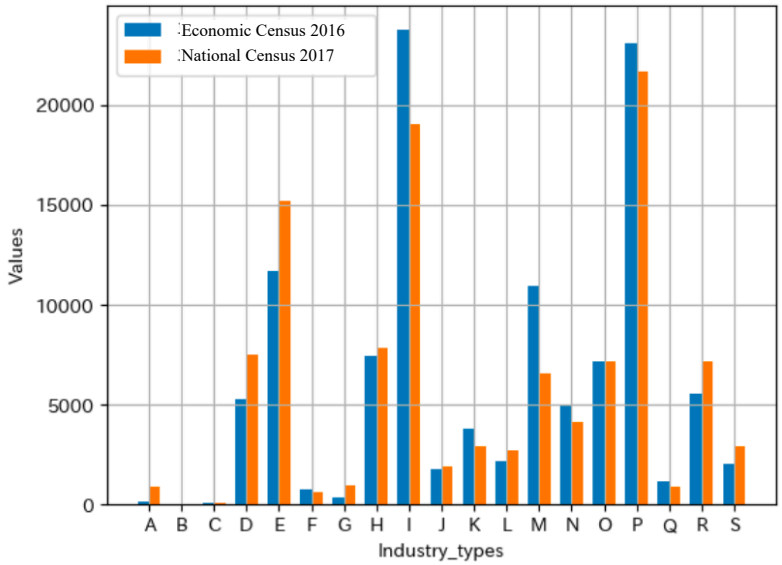

Figure 3. The number of workers by industry types in National Census in 2017 and Economic Census in 2016.

TABLE III. INDUSTRY TYPES IN NATIONAL CENSUS AND ECONOMIC CENSUS

Industry Type	
A) Agriculture & Forestry	K) Real Estate, Good Rental & Leasing
B) Fisheries	L) Scientific, Professional, & Technical Services
C) Mining	M) Accommodations & Eating and Drinking Serv.
D) Construction	N) Living and Personal & Amusement Services
E) Manufacturing	O) Education, Learning Support
F) Electricity, Gas, Heat Supply & Water	P) Medical, Health Care & Welfare
G) Information & Communications	Q) Compound Services
H) Transport & Postal Services	R) Other Services
I) Wholesale & Retail Trade	S) Government
J) Finance & Insurance	

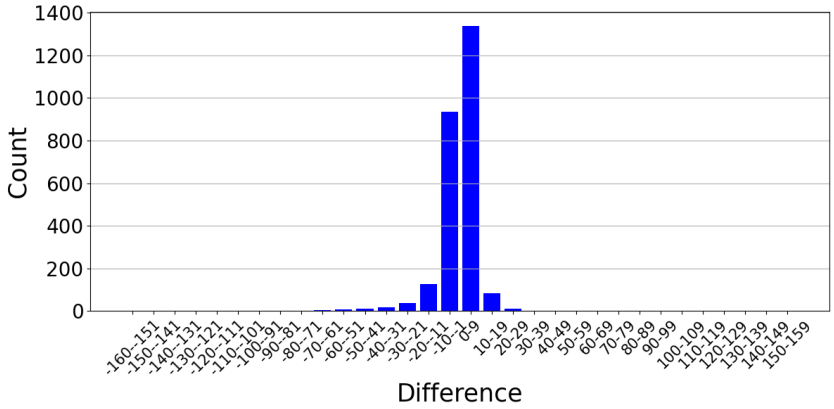

Figure 4. Address and attributes of household members (Small Area Map by Geoshape Local Government ID Dataset).

TABLE IV. THE NUMBER OF THE COMBINATIONS OF SMALL AREA AND INDUSTRY

Absolute Difference	The number of small areas by each industry
0-9	2,274
10-19	206
20-29	51
30-39	18
40-49	10
50 or more	19
Total	2,578

IV. PROPOSED METHOD

In this paper, we reassign the workers who are assigned to small areas that exceed the number of workers of statistics to small areas with a smaller number of workers assigned. After assigned by the former method [1], we apply our adjusting method to reassign the workers. We reassign the workers using the following procedure:

Step 1: Using the former method [1], assign workers to a small area in the target city.

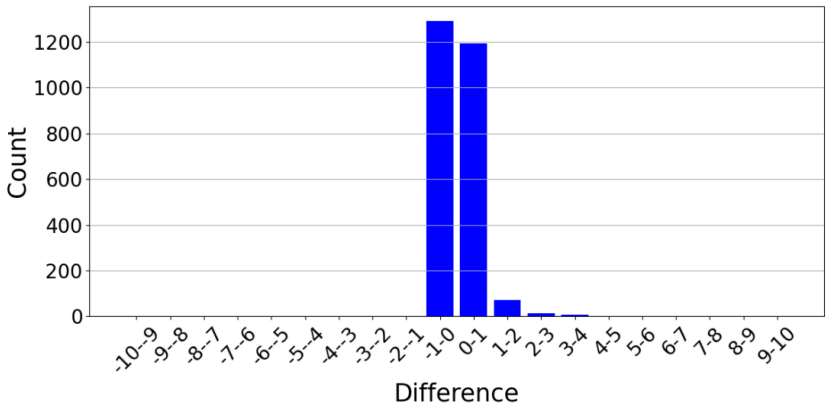

Figure 5. The difference in the number workers by small areas between the proposed algorithm and the adjusted statistics.

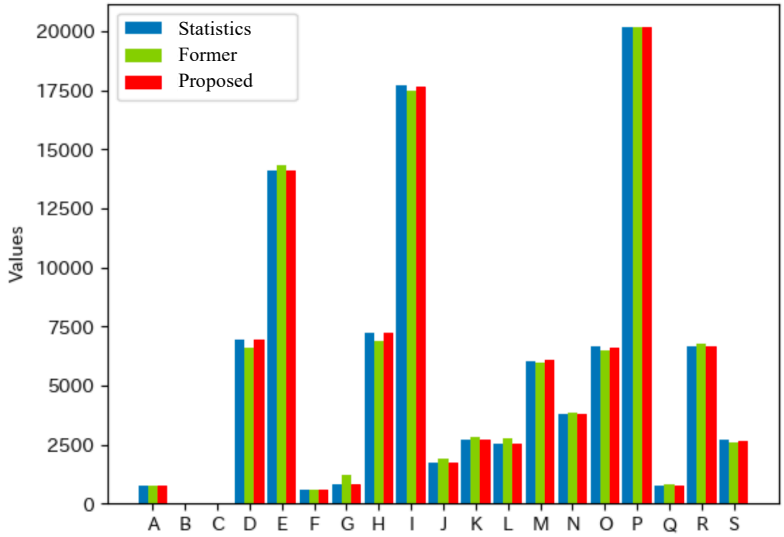

Figure 6. The number of workers in each industry by the former and the proposed methods.

TABLE V. THE NUMBER OF THE COMBINATIONS OF SMALL AREA AND INDUSTRY

Absolute Difference	The number of small areas and industry
0-1	2,486
1-2	71
2-3	13
3-4	6
4-5	2
6 or more	0
Total	2,578

Step 2: Find a combination of Small area s and Industry i with $Diff_{s,i} \leq -1$, then sort them in ascending order. When there is the same value of $Diff_{s,i}$, calculate the relative difference of $(E_{s,i} - S_{s,i})/E_{s,i}$, and find the larger one.

Step 3: Find a combination of Small area s and Industry i with $Diff_{s,i} \geq 1$. Let us denote a group of these combinations as *Exceed*.

Step 4: Find the combination of Small area s and Industry i with the smallest value of $Diff_{s,i}$ (the small area that has the largest shortage of the worker). From the small areas in *Exceed*, find workers to assign them to the small areas with the shortage of workers. According to the number of workers that exceed, randomly select a worker from those small areas and assign them to the target small area.

Step 5: Repeat Steps 2 to 4, until all combinations with $Diff_{s,i} \leq -1$ in Step 2 has been solved.

V. EXPERIMENTAL RESULTS

Fig. 5 shows a histogram of the number of combinations of Small area s and Industry i with $Diff_{s,i}$ using the proposed method. **TABLE V** shows the number of combinations of Small area s and Industry i in the absolute difference. From these figure and table, we can see that the difference between the number of workers assigned by the proposed algorithm

and the number of workers adjusted by Equation (1) by each small area becomes small. Since all the combinations come to the category 0-10, we make smaller categories 0-1, 1-2, 2-3, ..., 5-6, 6 or more. Almost all the combinations become less than 3 in $Diff_{s,i}$. By calculating (the total number of difference) /(the number of combinations of Small area *s* and Industry *i*), we can find that the difference of 0.48 workers averagely in each Small area *s* and Industry *i*, that is the difference per area and industry becomes from 4.22 to 0.48 by the proposed method.

Fig. 6 shows the number of workers by industries by the former method [1] and the proposed method. We can see that the proposed method improves the difference of the number of workers in each industry. Using the proposed method, we can reassign the workers in the small areas where there are no workers in the Survey. This operation can improve the number of workers in each industry.

VI. Conclusion

In this paper, we propose a method that reassigns workers in the exceeding small areas to the small area with the shortage of the workers. In the former method, there is a stochastic procedure to assign workers in small areas. That is the reason why there are several small areas with excess or shortage of workers. Using the proposed method, we can have a so-called "daytime population distribution" that is more compatible to the survey results.

As for further challenges of the proposed method, we should consider the sensitivity of the proposed method to apply. In this paper, we apply our method to a municipality with more than 200,000 population. Because municipalities with less than 200,000 population have less information about the workplace survey results, it is important to examine the reassignment of the proposed method. Since workers who work in a target municipality live in multiple municipalities, some of them come from municipalities with less than 200,000 population. That causes some difference in the number of workers in a specific industry. These challenges are left for further research.

References

[1] T. Murata, D. Iwase, T. Harada, "Workplace assignment to workers in synthetic populations in Japan," *IEEE Transactions on Computational Social Systems*, vol.10, no.4, pp.1914–1923, 2023.

[2] M. M. M. Abdel-Aal, "Calibrating a trip distribution gravity model stratified by the trip purposes for the city of Alexandria,f *Alexandria Engineering Journal* (Elsevior), vol. 53, pp. 677–689, 2014.

[3] N. Fournier, E. Christofa, A. P. Akkinepally, C. L. Azevedo, "Integrated population synthesis and workplace assignment using an efficient optimization-based person-household matching method," *Transportation*, vol. 48, pp. 1061–1087, 2021.

[4] P. Ye, F. Zhu, S. Sabri, F. -Y. Wang, "Consistent Population Synthesis with Multi-Social Relationships Based on Tensor Decomposition," *IEEE Trans. On Intelligent Transportation Systems,* vol. 21, no. 5, pp. 2180–2189, 2020.

[5] S. Agriesti, C. Roncoli, B. Nahmias-Biran, Assignment of a synthetic population for activity-based modeling employing publicly available data, *ISPRS International Journal of Geo-Information*, vol.11, 148, 26 pages, 2022.

[6] T. Murata, D. Masui, T. Harada, "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.

[7] T. Harada, T. Murata, "Projecting households of synthetic population on buildings using fundamental geospatial data," *SICE Journal of Control, Measurement, and System Integration*, vol. 10, no. 6, pp. 505–512, 2017.

[8] T. Murata, S. Sugiura, T. Harada, T., "Income allocation to each worker in synthetic populations using basic survey on wage structure, " *Proceedings of IEEE Symposium Series on Computational Intelligence* (Hawaii, USA), pp. 471–476, 2017.