Title	Configuring Agents' Attributes with Simulated Annealing
Author(s)	Hara, Shunsuke; Kita, Hajime; Ikeda, Kokolo; Susukita, Masahiro
Citation	Agent-Based Social Systems, 10: 45-59
Issue Date	2013-02-18
Type	Journal Article
Text version	author
URL	http://hdl.handle.net/10119/12803
Rights	This is the author-created version of Springer, Shunsuke Hara, Hajime Kita, Kokolo Ikeda, Masahiro Susukita, Agent-Based Approaches in Economic and Social Complex Systems VII: Post-Proceedings of The AESCS International Workshop 2012, Agent-Based Social Systems Volume 10, 2013, pp 45-59. The original publication is available at www.springerlink.com, http://dx.doi.org/10.1007/978-4-431-54279-7_4
Description	Agent-Based Approaches in Economic and Social Complex Systems VII: Post-Proceedings of The AESCS International Workshop 2012

Configuring Agents’ Attributes with Simulated Annealing

Shunsuke Hara, Hajime Kita, Kokolo Ikeda and Masahiro Susukita

Abstract In agent-based social simulation aiming at quantitative analysis of real situation faces a problem of deciding agents’ attributes. It has to match the existing data such as social statistics. However, the number of available data is much smaller than the vast degrees of freedom of agents, and therefore agents’ attributes can’t be decided uniquely. For this problem, the authors propose a formulation of attribute decision problem as a constrained optimization. In this paper, we show estimation of future population with an individual-based model as an example of the proposed method.

Key words: Agent-Based Simulation, Micro Simulation, Population Dynamics, Constrained Optimization, Simulated Annealing

1 Introduction

Computer simulation is utilized in various fields as the experimental methods for understanding, predicting and verifying complicated phenomena. As a simulation method for social science, the system dynamics (SD), a method that formulates behavior of the system to be studied as differential equations of the macro-level variables, is widely used. While SD enables to observe the behavior of whole system, it is difficult to study the society as micro-level behaviors such as actions of an individual people with this method. For solving this problem, the study of the agent-based simulation(ABS) is progressing with improvement in computational

Shunsuke Hara, Hajime Kita
Kyoto University

Kokolo Ikeda
Japan Advanced Institute of Science and Technology(JAIST)

Masahiro Susukita
Kansai Electric Power Co. Inc.

power[1], and object-oriented programming and artificial intelligence technology. With ABS, the social phenomenon is simulated in bottom-up manner by describing action of individuals as software agents and calculating the interaction of people or organizations.

However, if we try to quantitative analysis of actual society with the ABS, we face a difficult of deciding agents' parameters. The number of available data such as social statistics is much smaller than the vast degrees of freedom of the agents, and therefore agents' attributes can't be decided uniquely.

For this problem, the authors propose a formulation of attribute decision problem as a constrained optimization that is to find agents' attributes whose aggregated values meet social statistics. With simulated annealing, we sample agents attributes randomly from the feasible region of the optimization problem. Then with the sampled attributes, multiple runs of simulation are carried out and the results are studied as ensemble. In this paper, we show estimation of future population with an individual-based model as an example of the proposed method.

2 Population Problem of Aging Society

Japan, as well as other developed countries, is facing problems of aging and decreasing of population. It is a nationwide problems, but examining the problem regionally, the problem gets more serious. For example, population may decrease beyond the limit of viability of rural areas. In Japan, we often refer the problems as "low birthrate and longevity", the future population composition in Japan has the following characteristics:

- The birthrate is less than the level which can maintain population, and it is thought that population will decrease sharply.
- Because of "Baby boomer" generation's retirement, elderly people are increasing in number rapidly.
- On the other hand, youth population decreases remarkably by the double factors of decline of the population of the woman in a childbirth term, and the low birthrate.
- Aging of the residents of the bedroom suburb near big cities, where was developed at the postwar period, becomes serious.
- In the intermediate and mountainous area, aging and decreasing of population will go beyond the viability limit of community.

It is expected that these trends of population bring serious influence to the socioeconomic activity in Japan. For example,

- The ratio of elderly people to be supported to a productive person increases sharply.
- We have to expand quickly the medical treatment and the nursing home for elderly people.

- Because of decrease in population, the domestic market shrinks and the economic activity stagnates.
- The economic efficiency of maintaining lifelines to rural areas will be decrease.

Considering the aforesaid problem, the long-term strategic investment of infrastructures, such as electric power supply, medical institutions, and schools have to be done carefully with policy evaluation based of prediction of future household composition and population-by-age composition.

3 Individual-based Models of Population Dynamics and Their Problems

3.1 Macroscopic Model of Population Dynamics

The Cohort-Component Method (CCM) is a method used standardly as a technique of the population estimation used for population study.

In this method, numbers of individual age groups called cohorts are taken as variables, and evolution of population is modeled by difference equations of these variables considering birth, death, movement, and aging. That is, the model describes behavior of macroscopic variables of population. This method is useful for analysis of natural factors of population change, and model parameters can be estimated rather easily from the statistic data obtained by population census. On the other hand, it is hard to consider the effects of concrete policies which affect behaviors of individuals into estimation model.

3.2 Individual-based Models of Population Dynamics

Another approach of modeling population dynamics is population estimation by microscopic models such as agent-based model and microsimulation[2]. We call these models as Individual-based Models (IBM). In these models, behaviors of each individual such as birth, death, movement are described directly. IBMs have advantages and disadvantages contrary to the Cohort-Component Method. That is, to take effect of a concrete policy into consideration is rather easy because the model describes the microscopic behaviors of individuals. Disadvantages of the IBMs are 1) scale of the model, and 2) estimation of parameters. Because IBMs models each individual directly, the model size gets larger than the CCM which uses aggregated values as model variables. However, memory size and processing speed of current computer enable simulation of large scale IBMs, and this problem is not serious. More serious problem is estimation of model parameters. In IBMs, we have to decide various attributes of individuals used in the model. One approach is to use raw data of statistical survey. Microsimulation model INAHSIM by Inagaki[3] used this approach.

Agent-based simulation by Yamamoto[4] treats small village, and they could take similar approach. However, using survey data is restricted by availability of the data. If we have to use only statistics publicly open, we have to decide model parameters whose degree of freedom is much larger than the statistical data.

3.3 Related Study

INAHSIM uses microsimulation, and constructed in the 1980s as a future estimation model of households in Japan. Then, much improvements are made and it has become a tool usable as a synthetic tool for a policy simulation in the latest version.

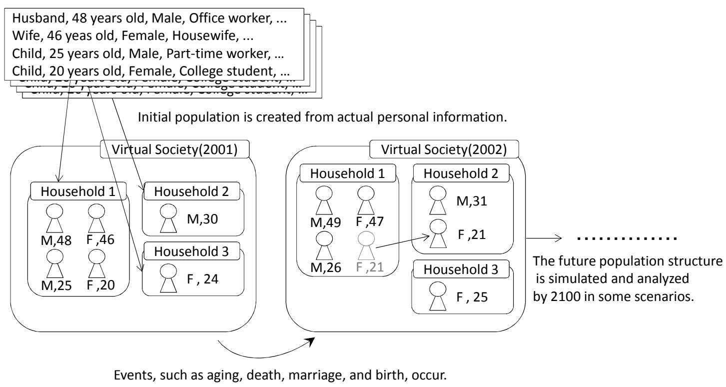

Fig. 1 Structure of INAHSIM

First, initial population is created as 1/1000 of Japanese society using the individual data of Comprehensive Survey of Living Conditions(2001), which is conducted by the Ministry of Health, Labour and Welfare. Information, including an individual birth year, sex, household composition, marital status, health condition, an employment state, earned income, pension benefit, lifetime income, and the history of these attributes, etc., is taken from the survey data and used for initial values of the model. Nine events, such as marriage, birth, death, divorce, international migration, and the change of health condition and an employment state, are generated according to each transition probability and dynamics of population is simulated with one year as the simulation step. Transition probabilities are presumed from the statistics related to each event, and they are assumed to be mostly constant during the time horizon of simulation. Structure of INAHSIM is illustrated in Fig. 1.

4 Determination of Agents' Attribute Value with Constrained Optimization

For quantitative analysis of population, agents' attributes in IBMs have to be decided with corresponding values of actual social surveys. However, degree of freedom of agents' attributes is much larger than the numbers of the actual statistic values that have to be considered. Hence, we can't decide agents attributes uniquely. To cope with this problem, we propose the following method:

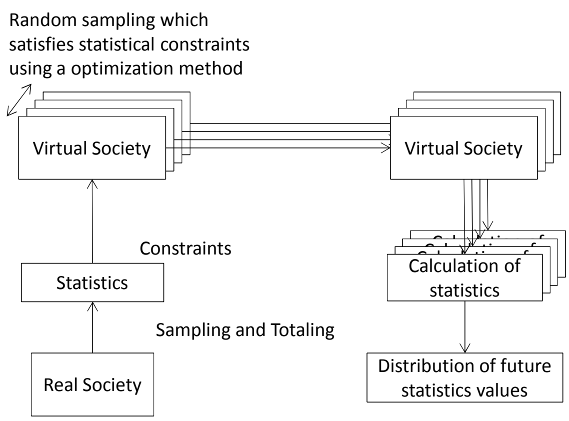

Fig. 2 Random sampling with constraint conditions and ensemble simulation

4.1 Setup Constraint Conditions

- Statistic values of social surveys to be considered are taken as the constraint conditions which should be fulfilled when deciding agents' attribute values.
- We also introduce constraint conditions that forbid agents' attribute to take invalid values from institutional point of view.
- Agents' attribute values are sampled randomly in the feasible region which fulfills the above-mentioned constraint conditions.

4.2 Random Sampling with Simulated Annealing

In order to constitute virtual society following the above setup, the problem is reformulated as an optimization problem considering constraint conditions as penalty functions to find solution in the feasible region of the constrained problem. Then we apply Simulated Annealing(SA)[5] for optimization. Theoretically, with sufficiently slow annealing schedule, SA samples solution uniformly and randomly from the feasible region of the problem. Although SA is usually regarded as an optimization technique, we used it as a random sampling method under constraint conditions.

4.3 Ensemble Simulation

Since the degree of freedom of the model is much larger than the considered conditions, configuration of agents' attributes is not decided uniquely. Hence, we repeat various runs of simulation with different parameter sets sampled by SA. That is, we carried out simulation as ensemble. Then the simulation results such as population size in future are examined looking at their distribution.

5 Individual-based Simulation Model of Population Dynamics

5.1 Outline

This model simulates regional evolution of population by aging, marriage, divorce, birth, death, and migration with one year as simulation cycle. Whole a country is consists of several area, and each individual belong to a household which locates in an area. Structure of the model is illustrated in Fig. 3.

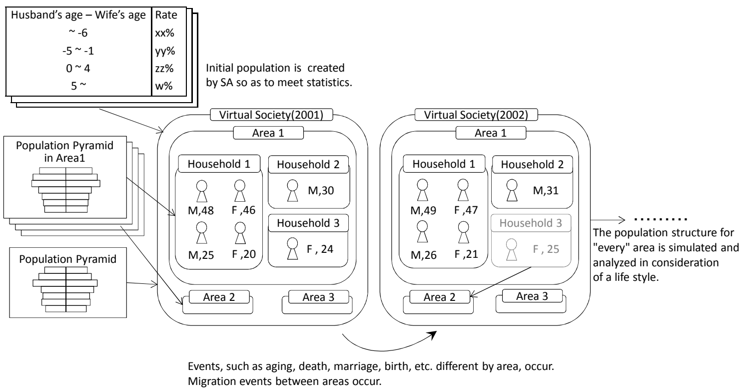

Fig. 3 Simulation model

After formation of an initial population, the following 6 events are calculated in each simulation cycle:

Aging Event : Since the simulation cycle is one year, agents grow older by 1 year at a time.
Marriage Event : Some pairs of single man aged over 18 and single woman aged over 16 will be in marriage state. For simplicity, we assume that the woman in marriage leaves her household and join the household where men belongs.
Divorce Event : It occurs to some household having a married couple. Currently we assume that the husband remains in the same household, and the wife generates a new household in the same area.
Birth Event : In some of household in which husband and wife exist, this event occurs in the probability according to the agents' attribute values, and add a 0-year-old agent to the household.
Death Event : An agent dies according to the probability specific to age and sex.
Migration Event : Some households move between the areas. In addition, division and integration of households are also considered. Assumptions on household formation made in the Marriage and Divorce Events are adjusted in this event.

Currently, we restrict our simulation to 1-region model, and the migration event is implemented only considering household division and integration. In the following formation of each step is explained.

5.2 Initialization of Population

Initial attributes of agents and households are decided by optimization technique proposed by [6]. In this research, Ikeda et al. formulated generation of the individual data which suits two or more amounts of totals as an optimization problem, and generate the agents group using the above-mentioned SA method. They constituted the households by the agents and has succeeded in generating the households of various life styles, such as husband and wife, children, and parents.

5.3 Marriage Event

Currently, two agents married regardless of area for simplicity.

Table 1 The marriage rate according to first-marriage couple's age difference (1995)

age difference	≥ 7	6	5	4	3	3	1	0	−1	−2	−3	≤ −4
ratio	10.9	4.4	5.9	7.8	9.7	11.6	14.3	17.6	8.1	3.8	2.3	3.5

5.3.1 Modeling of a Marriage with Constrained Optimization.

Two agents' group are made as decision variables, some statistical indicators are used to set constraint conditions, and we try to obtain a feasible solution, using the SA. Marriage occurs between a single male agent whose age is older than or equal to 18, and a female aged whose age is older than or equal to 16 considering Japanese law on marriage.

The following Constraints 1 and 2 are to fit to the male no-spouse marriage rates and the female non-spouse marriage rates in the Special Report on Vital Statistics(1995) by Ministry of Health, Labour and Welfare exhibits, respectively, and Constraint 3 is to meet the statistics about the marriage of Vital Statistics[7].

Constraint 1 : The male marriage rate according to age (with no distinction of the rate of the first marriage, and a second marriage rate)
Constraint 2 : The female marriage rate according to age (with no distinction of the rate of the first marriage, and a second marriage rate)
Constraint 3 : Distribution of differences between the husband's age and his wife's age (male agent age - female agent age) as shown in Table 1

5.3.2 Performance Function

The performance function used in a marriage event is designed like the performance function used when making initial population.

Marital relation $\mathbf{m} = \{(M_i, F_i)\}_i \in M$ is decided to optimize the performance faction where (M_i, F_i) expresses male agent M_i and female agent F_i in marital relation, and the range of i expresses $i = 1...N$ where N is the number of marriages to make.

N is calculated by

$$N = (\text{the number of alive agents}) * (\text{marriage rate})/1000.0 \tag{1}$$

and for the "marriage rate", we uses the marriage rate (number per 1,000 pairs of population) of vital statistics to the Ministry of Health, Labour and Welfare(1995).

The functions expressing the difference of $\mathbf{m}$ and each of three statistics data listed before are defined as $f_1(\mathbf{m})$, $f_2(\mathbf{m})$ and $f_3(\mathbf{m})$. Next, they are integrated into a single performance function

- Give a parameter to be evaluated ($\mathbf{m}$) and List l like Table 1 that describes the statistics. Let G be the number of items in List l.
- For item j in $j = 1...G$. As an example, an error function is designed for Table1.
  - Condition X on List l is expressed as X_j, and the actual rate shown as a statistics value is expressed as r_j.
  - Let c_j be the "the number of agents or groups which fills condition X_j" in$\mathbf{m}$.
  - Using the marriage number N, $N \cdot r_j$ is the actual numerical value as "the number of agents or groups which fills condition X_j".

- An error function is defined as $f_l(\mathbf{m}) = \frac{4}{G} \cdot \sum_{j=1}^{G} (c_j - m_j \cdot r_j)^2$.

It tries to minimize the square sum of the difference of “the real number of agents and groups” taken from statistics, and “the number of agents and groups”. Since “the real number of agents and groups” is not necessarily an integer, difference around 0.5 should be allowed. If this difference is 0.5 in all items $f(A)$ becomes 1.0. That is, if this index is about 1, the generated marriage agents are assumed to be in agreement with distribution of the statistical data.

By summing up all the $f_l(\mathbf{m})$, we obtain the optimization problem to decide marriage agents as follows:

$$\min_{\mathbf{m} \in M} \sum_{l=1}^{3} f_l(\mathbf{m}) \tag{2}$$

5.4 Divorce Event

In treating divorce, we have to decide not only husband and wife pairs to be divorce, but also belonging of their childes.

5.4.1 Modeling of a Divorce as Constrained Optimization

Decision variables of this event are the agent playing a role of the husband of the household, the agent playing a role of the wife, and how the parental authority of the child(ren) dependent on the husband and wife is divided. Some statistical indicators are made constraint conditions, and we try to obtain feasible solution using the SA method as an optimization problem which is expressed with the penalty function.

The divorce event occurs at the household in which husband and wife exist. Moreover, it must be taken into consideration which has the parental authority of the child(ren) dependent on the husband and wife. The following three statistical data is used as constraint conditions from ”Statistics about Divorce” of Vital Statistics Special Report of the Ministry of Health, Labour and Welfare[8].

Constraint 1 : A husband's divorce rate according to age group
Constraint 2 : A wife's divorce rate according to age group
Constraint 3 : The divorce rate classified by person which has parental authority
Since “the divorce number classified by person which has parental authority” is opened to the public, it converts and uses in percentage.

It is shown in Table 2 as an example.

Table 2 A husband's divorce rate according to age group(1995)

Husband's age group	~19	20 ~ 24	25 ~ 29	30 ~ 34	35 ~ 39
Rate	40.67	34.76	18.19	10.66	6.85
Husband's age group	40 ~ 44	45 ~ 49	50 ~ 54	55 ~ 59	
Rate	5.10	3.89	2.62	1.52	

5.4.2 Performance Function

The agent playing a role of a husband in an object household and the agent playing a role of a wife are expressed H_i and W_i . Moreover, which agent has parental authority is included in decision variables as C_i. The following cases can be considered.

1. Divorce in case there is no child who has parental authority
2. Divorce in case a wife has all the children's parental authority
3. Divorce in case husband and wife divide parental authority mutually
4. Divorce in case a husband has all the children's parental authority

$\mathbf{d} = \{(H_i, W_i, C_i)\}_i \in D$ is expressed by the above decision variables. In a similar way with the marriage event, penalty functions $f_1(\mathbf{d})$, $f_2(\mathbf{d})$ and $f_3(\mathbf{d})$ corresponding to the aforesaid constraints are formulated. Then it is integrated as a single performance function:

$$\min_{\mathbf{d} \in D} \sum_{l=1}^{3} f_l(\mathbf{d}) \tag{3}$$

5.5 Birth Event

In the birth event, the model is not formulated as an optimization problem but a birth event occurs to each agent playing a role of wife in the probability according to her age. Table 3 of the statistics about birth of vital statistics special report [8] which the Ministry of Health, Labour and Welfare exhibits is used.

Table 3 Age-specific marital fertility rate(1995)

age group	15-19	20-24	25-29	30-34	35-39	40-44	45-49
rate*	508.0	234.1	128.2	37.7	6.9	0.7	0.0

*Number of birth per 1000 females who have a husband.

5.6 Death Event

Using Table 4 of death of the vital statistics which the Ministry of Health, Labour and Welfare exhibits like a birth event, a death event occurs in the probability according to the sex and age to each agent. Table 4 is the used data.

Table 4 Mortality rate according to sex and age group (partly omitted)

Age Group	0-4	5-9	10-14	15-19	20-24	⋯	80-84	85-89	90-94	95-99	100-
Male	6.5	0.2	0.2	0.6	0.7	⋯	94.8	156.5	267.3	500.0	800.0
Female	5.4	0.2	0.1	0.2	0.3	⋯	54.3	100.7	195.7	400.0	720.0

Number of death per 1000 people.

5.7 Migration Event

Currently, we treat only adjustment of household formation bias made by the marriage and divorce events through division and integration of households. It also uses the constrained optimization method. That is, households after previous events are adjusted so as to meet the statistics of households in the previous term.

6 Experiment

This section shows the results of experiment. It is carried out to examine the behavior of ensemble simulation with sampling by the simulated annealing. First, initial population is created based on the statistical data in 1995. The whole country of Japan is considered as one area model. We use 5000 households and 11992 agents are created by the optimization in the initialization stage. The evolution of population is simulates 10 times with the model using different initial population and random number series.

The results of simulation by changing the random number series 10 times are shown as Figure 4 through Figure 9. It shows population-by-age for male and female at years of 1995, 2005, 2015. In figures showing population in 1995 and 2005, actual statistical data are also plotted in dashed lines for reference. In optimizations of the events modeled as constrained optimization, value of error function less than 1 is obtained in each run.

Looking at simulation results in 2005, we find the population profiles obtained by simulation appear near the actual statistics. Distribution of populations by multiple runs show the effect of random sampling with constrained optimization. Fluctuation of population in ensemble simulation shows that

- It enlarges according the time evolution.
- Fluctuation is not so large in age ranges born before simulation time window.
- Contrary to this, distribution spread wider in age ranges born in simulation.

7 Conclusion

In this research, we propose a method to configure large agents attributes under small number of constraints given as social statistics. Our method is to formulated the configuration problem as an optimization problem taking the the constraints as penalty functions, and to solve them with Simulated Annealing. We applied this method to individual-based model of population dynamics. Simulation results shows the fluctuation of simulation results can be assessed by ensemble simulation. As future study, we are also planning to expand the model to multi-area one by implementing the inter region migration.

References

1. N. Gilbert and K. G. Troitzch (2005) *Simulation for the Social Scientist, 2nd Ed.* Open University Press.
2. F.C. Billari et al. (2003) Introduction: Agent-based computational demography. *F. C. Billari, Alexia Prskawetz eds., Agent-Based Computational Demography: Using Simulation to Improve Our Understanding of Demographic Behaviour*
3. S. Inagaki (2007) Future socio-demographic population structure of japan: Projections by a dynamic microsimulation model (inahsim) (japanese). *Japan Statistical Association*
4. K. Yamamoto, Y. Koyama, and H. Deguchi (2005) Demographic simulation of depopulation in the rural area using soars (in japanese). *Technical Committee on System Engineering*
5. Emile Aarts and Jan Korst (1989) *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing*. Wiley
6. K. Ikeda, H. Kita, and M. Susukita (2010) Indivisual data for regional demographic simulations (in japanese). *Technical Committee on System Engineering*
7. Labour Ministry of Health and Welfare. Sumamry of vital statistics (in japanese). http://www.mhlw.go.jp/toukei/saikin/hw/jinkou/suii00/marr6.html
8. Labour Ministry of Health and Welfare. Specified report of vital statistics in fy2006 (in japanese). http://www.mhlw.go.jp/toukei/saikin/hw/jinkou/tokusyu/konin06/konin06-6.html

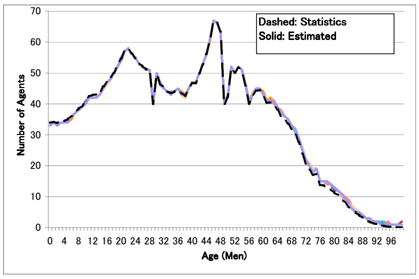

Fig. 4 Initial male population distribution (1995)

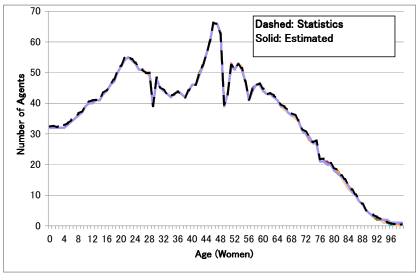

Fig. 5 Initial female population distribution (1995)

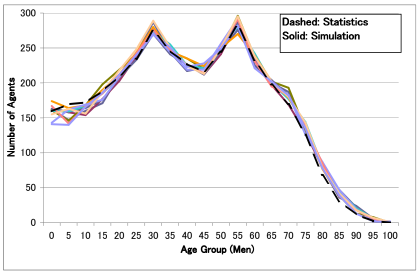

Fig. 6 The simulated result of male population distribution by 5-year age group(2005)

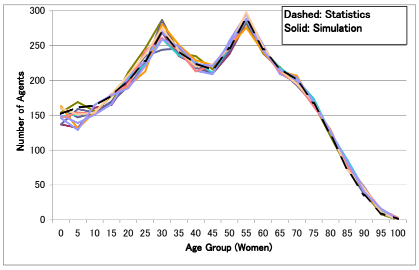

Fig. 7 The simulated result of female population distribution by 5-year age group(2005)

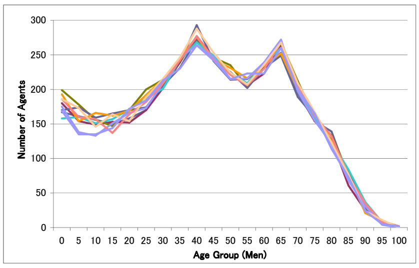

Fig. 8 The simulated result of male population distribution by 5-year age group(2015)

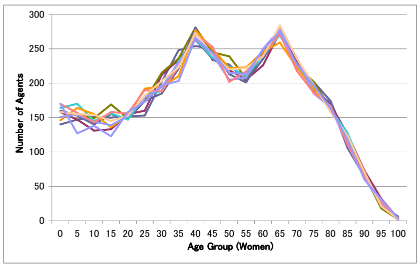

Fig. 9 The simulated result of female population distribution by 5-year age group(2015)