Agent-Based Simulation of School Choice in Bandung, Indonesia: The Emergence of Enrollment Pattern Trough Individual Preferences

Dhanan Sarwo Utomo, Utomo Sarjono Putro and Pri Hermawan
School of Business and Management, Institut Teknologi Bandung, Indonesia

ABSTRACT

This study is motivated by the reality that school choice programs that is currently implemented in Bandung that, always resulting student deficit (lack of student) in some schools. In this study, a mechanism that can describe how the enrollment pattern in a school choice program emerge as a result of individual preferences of the prospective students, is constructed. Using computer simulation, virtual experiments are conducted. In these experiments, the enrollment patterns and the number of student deficit that were resulted by various school choice program configurations are analyzed. Based on the experiment results, modification of the current program that can minimize the number of student deficit can be purposed.

Keywords: agent-based simulation, school choice, computer simulation

1. Introduction

1.1 School Choice Problem in Indonesia

As an impact of the implementation of Indonesian Republic Law No.32 2004, each local government has an obligation to design suitable education policy for their region. One of the obligations that should be met is to design a school choice program that is suitable to be implemented in their region. School choice program is a program that facilitates the parents and their children to select schools (Brown, 2004; Betts, Rice, Zau, Tang, & Koedel, 2006). In Indonesia this kind of program is also known as the acceptance of new students program (is abbreviated as PSB in Bahasa).

From year to year, school choice program in every region, especially for senior high school, is vary in the term of:

• The information technology that is used to manage applicants administration data. For example, central data base that can be accessed via internet is used to manage applicants administration data in Jogjakarta and DKI Jakarta (Rusqiyyati, 2008).
• They way the schools are classified or clustered. For example, in Bogor the schools are categorized or clustered based on their achievement from previous year experience while in Takengon the schools are clustered based on their geographic location (Muisman, 2003).
• The number of school to which the prospective students may enroll. For example, students in Yogyakarta may enroll to three schools while in Bandung, students is only allowed to enroll to two schools (Kompas, 2007).
• Admission criteria that is used to determine whether an applicant is accepted or not.
For instance, a number of regions in Indonesia use national examination score as the admission criteria. In another region like Bekasi a combination between national examination score and test score that held by each school is used as admission criteria (Pos Kota, 2008).

Education experts and practitioners in Indonesia argue that, there are two criteria of a good high school choice program, namely:

- Fair and free from corruption, collusion and nepotism practice (Junaidi, 2008; Suwarja, 2003; Vardhana, 2008).
- The incoming students are distributed evenly to all public and private school in that region (Suryadi, 2008; Fathoni, 2008).

Unfortunately, school choice programs that are currently applied in many cities in Indonesia still cannot meet these criteria. There are many schools that suffer from lack of students (Antara, 2008; Radar Cirebon, 2008; Banjarmasin Post, 2009). On the other hand, many prospective students were rejected because the schools to which they enroll have lack of space (Kompas, 2007; Sumatera Ekspres, 2007; Surya, 2009). In addition, there are gaps in the quality of incoming students between the favorite schools and less favorite schools (Siahaan, 2008).

To improve the performance of school choice program, the city governments continuously modify the program that is applied. These modifications are usually made based on the evaluation of the program’s performance in the previous year (Suara Merdeka, 2008; Radar Bogor, 2009). This mode is no other than an experiment that is conducted in real system. Experimentation in real system was very risky because a school choice program involves a complex social process. A complex social process consists of many non linear interactions among elements, in this case human (Gilbert, Agent-based Social Simulation: Dealing with Complexity, 2004). This non linear interaction is caused by human societies characteristic that can recognized and changed their behavior (adaptive) in order to respond to (Gilbert, Emergence in Social Simulation, 1995) the new implemented system. Because of this human society characteristic, the impact that will occur from the experiment will be very hard to predict (Agar, 2004).

1.1 Purpose of the Study

This study aims to purpose a mechanism that can describe how the enrollment pattern in a school choice program, can emerge from the individual preferences of prospective students. The case that is selected in this study is the school choice program in Bandung in 2008. Using computer simulation, experiments with various system configuration and agent’s characteristics are conducted. Specifically, these experiments aim to show the enrollment patterns and the number of student deficit that may occur. Based on the experiments results, school choice program that can minimize the number of student deficit can be purposed.

2. Modeling Process

2.1 Description of School Choice Program in Bandung

In 2008 there are 26 high schools in Bandung that participate in the school choice program. All of these schools are public high school. The rest of the high school and the private high schools have obligation to hold their own selection process.

26 high schools that participate in the school choice program are clustered into five. These clusters were made based on the school’s previous achievement. Prestigious high schools are placed in the first cluster and less favorite schools are placed in the last cluster. Each applicant may only choose two schools from different cluster. They have seven days to consider and submit their application to the selected school. Each applicant was able to access daily information about the number and
national examination score of all applicants that already submit their application.

There is no specific standard that have to be met by an applicant. The lowest national examination score of a student that was accepted in the previous year (known as passing grade) usually become guidance for the applicants to select a school. Schools with high passing grade usually interpreted as a prestigious schools, which is commonly avoided applicants with low national examination score.

2.2 Applicant Decision Making Process

There are three general steps that are usually taken by an applicant to decide schools to which he or she will apply. In the first step, an applicant will gather information about the schools (Tatar & Oktay, 2006). There are six kind of information that usually gathered by the applicants in this step, namely:

- Applicant’s residence location (Henrickson, 2003)
- Distance travelled to school (Henrickson, 2003; Tatar & Oktay, 2006)
- The number of application that have been sent, represent the number of competitors they will face.
- Applicant’s achievement represent by the national examination score.
- Applicant’s expectation about the school’s quality (Henrickson, 2003; Tatar & Oktay, 2006), that is based on the school’s previous achievements.
- The minimum qualifications that are accepted in that school (Tatar & Oktay, 2006).

Applicants have autonomy to determine the importance of each school attributes.

In reality, the applicants will not have complete information of all available schools. But, in order to simplify the problem at hand the simulation that will be constructed is based on the assumption that, all applicants (agents) are able to gather complete information about all school and, they will update this information regularly.

In the second step, the applicant will consider the school that is most appropriate for them. This process can be represents by the process of maximizing school’s aggregate benefit according to applicant (Belfield & Levin, 2002).

In the last step the applicant will consider whether he or she will be qualified in the selected schools (Belfield & Levin, 2002). This process can be considered as the process of comparing applicant’s national examination score to the school’s passing grade.

2.3 Purposed Agent-Based Simulation

The purposed simulation is constructed using SOARS (Spot Oriented Agent Role Simulator) that was developed by Deguchi Laboratory in Tokyo Institute of Technology. There are two types of object in SOARS, spot and agent.

In this model an agent represents an applicant in the school choice program. The total number of agents is 2850, that represent 6% of the total applicants in 2008. Agents are categorized into three types:

- Neutral agents: represent applicants who believe that the minimum score they should have to be qualified in a school can fully be described by the previous year passing grade. Therefore, they will apply to a school only if their score is higher than the previous year passing grade of that school.
- Pessimistic agents: represent the applicants who disbelieve that the minimum score they should have can be fully represented by the previous year passing grade. Therefore, they make some adjustment to anticipate in case the minimum the minimum score needed to be qualified increase.
- Optimistic agents: this type of agent also disbelieve that the minimum score they
should have can be fully represented by the previous year passing grade. But, they dare to make speculation and apply to the top school of their choice.

There are two types of spot that are defined in this model. The first is the home spot in which the participants stored all information during the simulation and the second, school spot to which the applicants enroll. The number of school spot is 26 (equal to the number of schools that participate in the school choice program in 2008) and the number of home spot is 2850 (equal to the number of agent). The next section describes mechanisms and attributes owned by each spot and agent.

1) Attributes of School Spot

The first attribute for the schools is \( x \) and \( y \) position in a grid of Bandung City. In order to assign schools position, Bandung city is divided into 6 \( \times \) 4 grids. Then, the \( x \) and \( y \) position for school spot is assigned based on the actual position of the given school in the real world.

The second attribute of the schools is passing grade score (PG). A passing grade score indicate the lowest national examination score that was accepted in a certain school in the previous year. In this study, passing grade score of each school is assign based on the result of the school choice program in 2007.

The third attributes of the school spot is capacity (\( C_j \)). School’s capacity indicates the number of student that can be admitted by a certain school. The capacity of each school spot is assigned based on the data in 2008.

The fourth attributes of school spot is image score (\( IM_j \)). Image score indicates the expected education quality of schools based on the perception of the agents. This attributes is assigned as a real number from 0 to 100. The higher image score, the higher the expected education quality of the given school. Although there is no research that directly measure image of the schools in Bandung city, some research in Indonesia revealed that there is a linear relationship between the passing grade score and the school’s expected education quality (Muisman, 2003; Purnawan, 2005; Priyanta, 2008; Rijanto, Hadi, & Relisa, 2008). Therefore, the following linear relationship is used to assign image score to each school spot in the simulation.

\[
IM_j = 7.651PG_j - 114.765
\] (1)

The fifth attribute of the school spot is the number of applicants (\( NA_j^t \)). This variable is an integer counter of the agents who enroll to school \( j \) at each time step. At the beginning of the simulation, this variable is initiated as zero.

The sixth attribute of the school is the cluster number (\( CL_j \)). This attribute is put into the simulation as integer based on the school’s cluster data in 2008.

The seventh attribute of the school is the minimum applicant’s score (\( MS_j^t \)). This variable indicates the lowest national examination score of all agents who enroll at school \( j \) at time \( t \). At the beginning of the simulation, this variable is initiated as zero.

![6 x 4 grid of Bandung city](image)

The eighth attribute of the school spot is the competitor score (\( CS_j^t \)). Competitor score is a real number that range from 0 to 100. This variable indicates the degree of competition that will be faced by agents if they enroll to school.
This variable is inserted in the simulation as a function of the capacity of school \( j \) and the number of applicants in school \( j \) in the previous day.

\[
CS_j^{t+1} = \frac{100 \times (C_j - NA_j^t)}{C_j}
\]  

(2)

The last attribute of the school spot are two arrays to store the indexes \( AI_j \) and national examination scores \( AS_j \) of the agents who have enroll to school \( j \). At the beginning of the simulation both of these arrays are initiated as .

2) Attributes of Home Spot

Just like the school spots, each home spot is also equipped with the \( x \) and \( y \) position. To assign the \( x \) and \( y \) position of home spots, the population densities at each point on the grid are calculated. The population densities are calculated by dividing the population in each point by the total population in Bandung. After that, the coordinate of each home spot is assigned using roulette wheel method in which, the population density serves as the probability.

The second attribute of the home spot is its distance to each school \( (r_{ij}) \). This attribute is calculated using Euclidean distance formula as the following:

\[
r_{ij} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}
\]  

(3)

Where, \( r_{ij} \) is the distance from home \( i \) to school \( j \), \( x_j \) is the position of school \( j \) in \( x \) axis, \( x_i \) is the position of home \( i \) in \( x \) axis, \( y_j \) is the position of school \( j \) in \( y \) axis and \( y_i \) is the position of home \( i \) in \( y \) axis.

The distance to each school is then converted into a score \( (D_{ij}) \) from 0 to 100. Score of 100 is assign if \( r_{ij} \) is equal to the diameter of the grid (7.211). In order to assign the distance score of each school the following equation is used.

\[
D_{ij} = -13.867 \times r_{ij} + 100
\]  

(4)

The third attribute of the home spot are two arrays to record index of the schools that are chosen by agent his or her first \((CH1_i)\) and second \((CH2_i)\) choice. At the beginning of the simulation, both of these arrays are initiated as .

The last attribute of the home spot are two arrays to record cluster code of the schools that are chosen by agent as his or her first \((CL1_i)\) and second \((CL2_i)\) choice. At the beginning of the simulation, both of these arrays are initiated as .

3) Agent’s Attributes

The first attribute of the agents is their national examination score \( (S_i) \). The minimum national examination score that should be achieved in order to be graduated from junior high school is 22 while, the maximum national examination score that can be achieved is 40. Therefore, in this study the national examination score of each agent is assigned as a random number from 22 to 40.

The second attribute of the agents is the weights for each school attribute namely, distance weight \( (WD_i) \), image weight \( (WIM_i) \) and competitor weight \( (WCS_i) \). All of these weights are assigned as a random number from 0 to 1 but, the total of all weights may not exceed 1. In order to fulfill this constraint, the following steps are taken.

\[
WD_i = \text{random}[0,1]
\]  

(5.a)

\[
\Delta = 1 - WD_i
\]  

(5.b)

\[
WIM_i = \text{random}[0,\Delta]
\]  

(5.c)

\[
WCS_i = \Delta - WIM_i
\]  

(5.d)
The third attribute for the agents is agent’s tolerance ($T_i$). Agent’s tolerance indicates agent’s boldness to speculate on the fluctuations of passing grade score that may occur. For neutral agents, tolerance value is set as zero.

$$T_i = 0 \quad (6.\text{a})$$

For optimistic agents, tolerance value is initiated as a negative random number with restriction that, the total of tolerance value and agent’s national examination score may not less than 22 (since there will be no school which passing grade is less than 22).

$$T_i = \text{random} \left[0, (22 - S_i)\right] \quad (6.\text{b})$$

For pessimistic agents, tolerance value is initiated as a positive random number with restriction that, the total of tolerance value and agent’s national examination score may not greater than 40 (since there will be no school which passing grade is greater than 40).

$$T_i = \text{random} \left[0, (40 - S_i)\right] \quad (6.\text{c})$$

4) Simulation Process

At every time step, from the 1st day until the 7th day, as long as $CL_{1i} = \emptyset$ or $CL_{2i} = \emptyset$, agent $i$ will visit all school. Agent $i$ then will record the passing grade score ($PG_j$), competitor score ($CS_j^i$), image score ($IM_j$), cluster code ($CL_j$) and the minimum applicant’s score ($MS_j^i$) from each school. These variables are then stored in agent $i$’s home.

In each home spot, agents calculated the aggregate benefit of all school using additive model (Goodwin & Wright, 2004).

$$AB_{ji} = WD_i D_{ij} + WIM_i IM_j + WCS_i CS_j^i \quad (7)$$

Where, $AB_{ji}$ is the aggregate benefit of school $j$ according to agent $i$.

The aggregate benefit scores of all schools are then sorted from the school with the highest aggregate benefit to the school with the lowest aggregate benefit. Agent $i$ then determines the school to which he or she will enroll. The evaluation process is started from the school with the highest aggregate benefit to the school with the lowest aggregate benefit. In this process agent $i$ will calculate the difference between his or her national examination score and the passing grade of school $j$ plus agent’s tolerance.

$$d_{1ij} = S_i - (PG_j + T_i) \quad (8)$$

If $d_{1ij} > 0$ then, agent $i$ will chose school $j$. If $d_{1ij} < 0$ then, agent $i$ will not chose school $j$ and, will continue the evaluation process to the next best school. If $d_{1ij} = 0$ then, agent $i$ evaluate whether he or she has better chance to be admitted than the applicant with the lowest national examination score at school $j$.

$$d_{2ij} = S_i - MS_j^i \quad (9)$$

If $d_{2ij} > 0$ then, agent $i$ will chose school $j$ else, agent $i$ will continue the evaluation process to the next best school.

Every time an agent chose a school, he or she will check whether he or she is able to enroll to the chosen school. If $CL_{1i} = \emptyset$ and $CL_{2i} = \emptyset$ then, agent $i$ will store the index of school $j$ in array of first choice index ($CH_{1i} = j$) and, the cluster code of school $j$ in the array of first choice cluster ($CL_{1i} = \{CL_j\}$). If $CL_{1i} \neq \emptyset$ and $CL_{2i} = \emptyset$ and $CL_j \notin CL_{1i}$ then, agent $i$ will store the index of school $j$ in array of second choice index ($CH_{2i} = \{f\}$) and, the cluster code of school $j$ in the array of second choice cluster ($CL_{2i} = \{CL_j\}$). If school $j$ cannot satisfy these conditions then, agent $i$ will chose new school.

After the evaluation process, agents will enroll to the schools whose index is stored in
array $CH1_i$ and $CH2_i$. Each agent who enrolls to school $j$ will increase the number of applicant at school $j$ ($NA_j^t$) by 1. Agent’s index and national examination score then will be stored in array of applicant index ($AI_j$) and array of applicant score ($AS_j$) at school $j$.

After agents enroll to the schools of their choice, each school then will update the competitor score ($CS_j^t$) and the minimum applicant’s score ($MS_j^t$) for the next iteration. The new competitor score in each school is calculated using (2).

To update the minimum applicant’s score in each school, both $AI_j$ and $AS_j$ are sorted based on the applicant’s score, from the highest score to the lowest score. If $NA_j^t \leq C_j$ then, the $MS_j^t$ is equal to the applicant’s score who is in the $NA_j^t$ rank. If $NA_j^t > C_j$ then the $MS_j^t$ is equal to the applicant’s score who is in the $C_j$ rank. After these processes are finished, the iteration counter is increased by 1.

At the end of the school choice program, both $AI_j$ and $AS_j$ in all schools are sorted based on the applicant’s score, from the highest score to the lowest score. If $NA_j^t \leq C_j$ then, the school $j$ will admit all agents who enroll to school $j$. If $NA_j^t > C_j$ then, the school $j$ will admit agents whose rank is less than or equal to $C_j$.

3. Experiment Process

In this study three experiments are conducted. The aim of the first experiment is to test the performance of the current program (five clusters with two choices) under various population variations. The second and third experiment aimed to test the performance of the school choice program using different number of cluster.

In each experiment, four kinds of scenarios are carried out. The first scenario involve only neutral agents, the second involve only pessimistic agents, the third involve only optimistic agents and in the last scenario all types of agents are involved with equal proportion.

In the first experiment observed that in the current school choice program, student deficit are always occurred in every scenario. The number of student deficit increases drastically in the second scenario, when all agents are pessimistic. This is happened because pessimistic agents tend to avoid competition, indicated by a low total number of applicants in each cluster.

In the second experiment, the number of cluster is reduced into three clusters. This program performs better in eliminating the number of student deficit in the second and fourth scenario. But, this program performs worse than the current school program in the first and third scenario.
In the third experiment, the number of cluster is increased into six clusters. This program performs better in minimizing student deficit in all scenarios. The deficit only occurred in the third scenario, with less number compare to the two previous programs.

From the experiment results it can be concluded that the current school choice program is very sensitive to the variation of population proportion and the number of cluster. The current school choice program is vulnerable in resulting high number of student deficit especially, when the number of pessimistic agents is high. In order to minimize the number of student deficit in each cluster, the decision maker in Bandung can increase the number of cluster that is used, from five clusters to six clusters.

4.2 Further Research

This study has several limitations that should be improved in the future. The first limitation is that the experiment result is still not yet validated externally. In the next study, we aim to compare the enrollment pattern and the number of student deficit in each cluster, to the data taken from the real world.
The second limitation of this study is that, the agents are only interact each other through the school spots. The agents are also assumed to have complete information about all schools attributes. Studies have revealed that in a school choice program, agents have incomplete information about the school’s attribute. Agents will rely on their social networks (for example: ex-schoolmate, parent’s co-worker, or family member) in gathering more information about the school attributes (Holme, 2002; Ramsay & Sanchez, 2006; Dougherty, et al., 2007; Dillon, 2008). In the next study we aim to improve the purposed mechanism, in order to facilitate these behaviors.

References


