# GIS BASED MODEL FOR ESTIMATING THE NUMBER OF BUS PASSENGERS

*Ryoko FUNATO, Master Student, Graduate School of Science & Engineering, Saitama University, funato@dp.civil.saitama-u.ac.jp*

*Kunihiro SAKAMOTO, Assistant Professor, Graduate School of Science & Engineering, Saitama University, sakamoto@dp.civil.saitama-u.ac.jp*

*Masaru YAJIMA, (President of Eagle Bus Co., Ltd.), PhD Candidate, Graduate School of Science & Engineering, Saitama University, yajima@dp.civil.saitama-u.ac.jp*

*Jyunichi YAMAGISHI, Area Market Structure Laboratory Co., Ltd., yamagisi@dp.civil.saitama-u.ac.jp*

*Hisashi KUBOTA, Professor, Graduate School of Science & Engineering, Saitama University, hisashi@dp.civil.saitama-u.ac.jp*

## ABSTRACT

This study develops the BPE model (Bus stop Passenger Estimate model) which calculates the number of passengers at each bus stop to examine the optimal bus stop position with combining existing data of GIS (Geographic Information System). As a case study, the BPE model was developed with the multiple regression analysis for a bus route in suburban area city, Hidaka city. The bus enterprise introduces a new system that enables us to accurately grasp the present number of bus users with a unit of bus stops. In addition the analysis uses other new system named AMS system (Area Market Simulator system), which is a developed function of GIS and helps us to make micromesh data. In the case two types of model are made; one is commute hour model and another is non-commute hour model. The BPE model is developed by using population, bus frequency, and distance from home to railway station. The both type models are significant. The BPE model forecasts the number of passenger for alternative routes. This BPE model will be helpful to the bus enterprises and the municipalities for the route maintenance.

*Keywords: public transportation, bus service, GIS*

# INTRODUCTION

## The present conditions of the Japanese bus businesses

In Japan, bus businesses are declining for a long term. The number of bus users is decreasing because people depend on private car. The situation is hard for bus enterprises.
Naturally, hereditarily or systematically Japanese bus enterprises have been played all roles in bus businesses; “plan”, “management”, and “service.” In late years a new system, cooperative system, is in the spotlight. For example, particularly in underpopulated areas, a municipality functions as planning and management, and an enterprise serves bus services. The consultation about the activation of the public transportation is increasing all over the country. But, the old system is kept in most areas.
It seems that the basic things, such as a bus route or frequency of bus services, were made in old days, and habitually the old regulations are used even now. Freedom of the competition is promoted by the revision of the law about the regulation of new entry in 2004. The bus enterprises are demanded to improve the management for the competition, in addition, for regaining bus users. However, it is difficult to suddenly improve the ingrained customs. It seems that they don't know the way of improvement.
The enterprises that manage deficit routes are mostly compensated for the administrative cost as subsidy from the local government. Because bus services can be distributed as a public works project, it is reasonable that municipalities subsidize bus enterprises that own deficit routes, but earlier than anything enterprises must make an effort to improve their bus service. Nevertheless, in the present state of affairs, the enterprise is not ready to improve the management. They need to know the present number of bus passenger, and to improve the basic bus services for increasing bus users.

## Goal

Few enterprises can definitely answer these questions; "Why the bus route is run over in the present route?" and "Why there is a bus stop at the present position?”; because a lot of bus routes have been taken over from the previous route all the time. This research aims for developing BPE models as one of tool which is used to improve bus services.
By using the BPE models, planners can compare more than one bus stop positions (Figure 1). The BPE models are basically based on population around bus stop. They can try to find out improved bus stop(s) and/or route(s), which should have more bus passengers. Here, the route is improved under the condition that "the both distances of the present route and the proposed route are at the same level". It can be beneficial for bus enterprises because they can count on much return. Furthermore, it also can be beneficial for the residents near the bus route areas because the chosen bus stops will be what are usable for many citizens.

## Demand prediction on bus services

There are a lot of studies about demand prediction for bus till now. For example, the questionnaire analysis is a typical method (for example YOSHIDA.I 2004, FUJII.S 2003). On

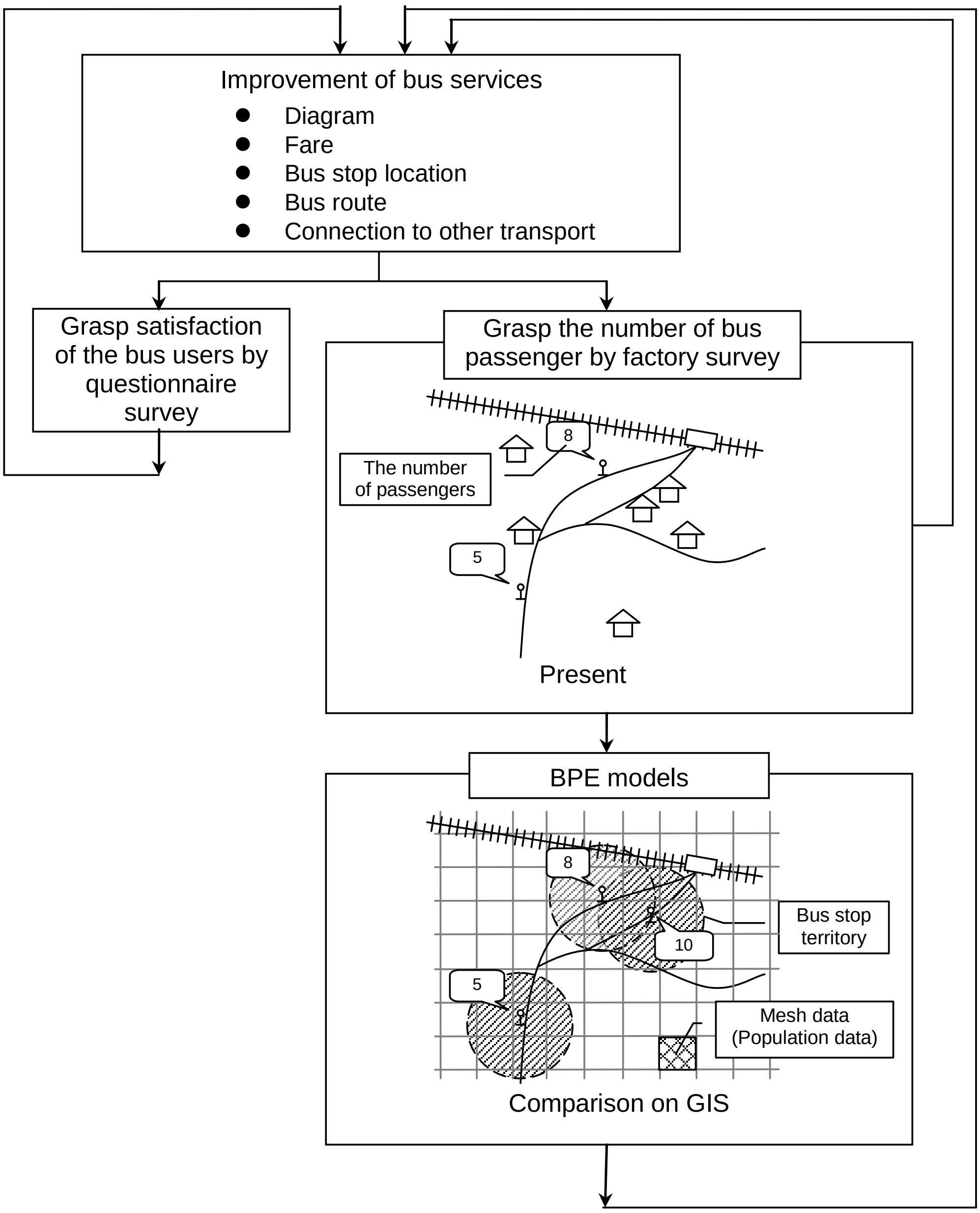


Figure 1 – Role of the BPE models

the other hand the study utilizes existing information such as population or a bus service level. Bus enterprises will feel comparatively easy to introduce the tool because of the use of existing data.

There are some studies to analyze the demand for bus with existing data, but they deal with a city bus or a minimum bus (for example SUGIO.K 2001). Because of a large population, it is thought that a city bus is under the condition that the businesses are apt to be in surplus. A minimum bus cannot but receive subsidy from municipalities. This study pays attention not to urban area or underpopulated area but to suburbs. Suburb is a delicate area. Bus enterprises are hard to predict whether they can get profit or not in suburb. But it can be said that in suburban area bus enterprises have a possibility to get rid of a deficit by an improvement of bus service.

And the existing studies evaluate the number of passengers with unit of a road or a route. This study handles a micro issue that is bus stop positions. The micro analysis needs further information. AMS made it possible to make micromesh easily and helps the analysis go along smoothly. Use of the result of the micro analysis improves the detailed bus services. The number of passengers will be increased not by big change of route but by small review of bus stop positions.

# DEVELOPING THE BPE MODELS

The BPE model is made by use of multiple regression analysis. Factors in the number of bus passengers are classified three types; the demographic and social economic factors, bus service level, and others. For example the demographic and social economic factors are household, population, school, hospital, station and so on. They are classified the origin factors, such as the number of the households, into the destination factors, such as shopping stores. Bus service characteristics are such as frequency, fare, bus stop, required time. Other factors are driver's license, private car, landform, and parking.
The use of two systems helps with the making of the BPE models. Data of the number of passengers is counted by sensor system, and micromesh data is made by using the AMS system. The relationship is shown by Figure 2.

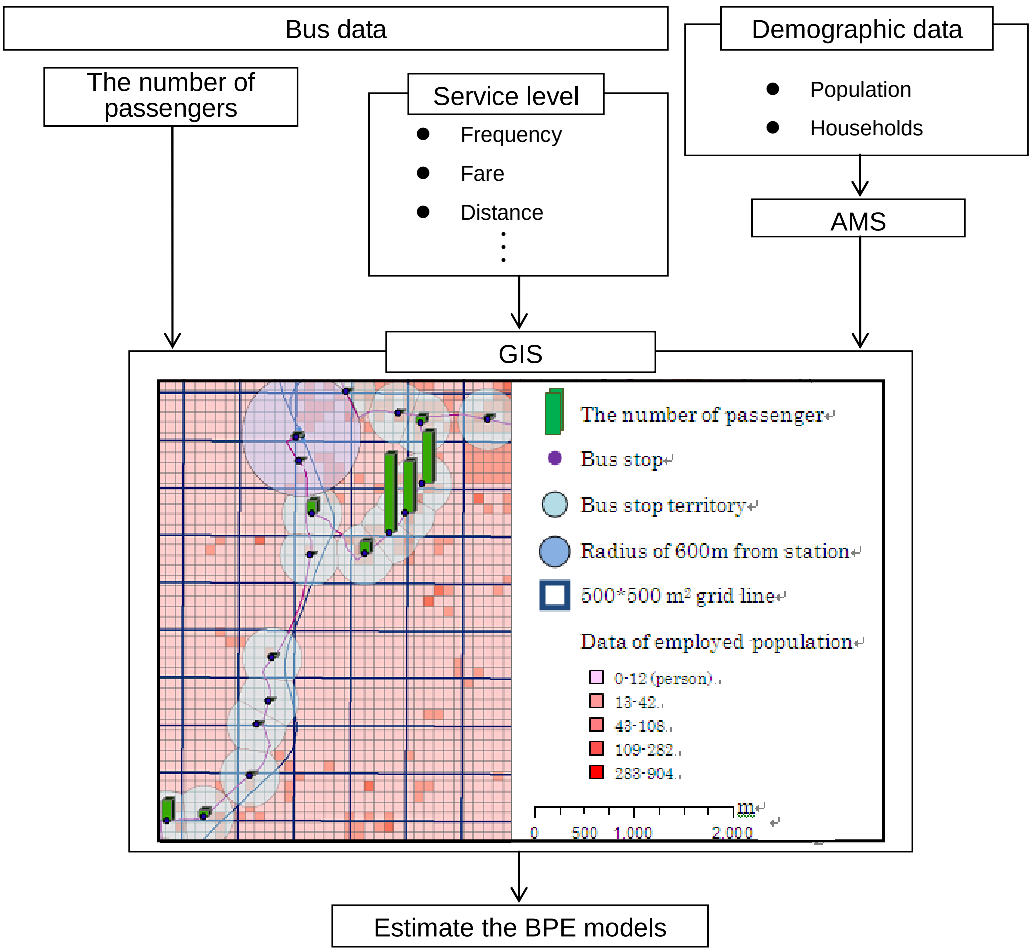

Figure 2 – Image of the study

### Micromesh data made by using AMS system

Populations within a bus stop territory are calculated as demographic and social economic factors. GIS, Arc View9 of ESRI Company, was used for calculate the partial population. The space analysis function is one of the characteristics of GIS. The populations are assigned based on the occupied area within the grid. The bus stop territories are defined the range that people can walk to a bus stop.
Digital map 25000 published by the Geographical Survey Institute is a usable data. The analysis should use grid size of 100*100 m2, which is 1/25 size of the usual grid size of 500*500 $m^2$. Figure 3 displays GIS map with grid data of 500*500 $m^2$ size and grid data of 100*100 $m^2$ size. The deeper the red colour of the grids, mean more people live there. We can understand that there are biases of population within 500*500 $m^2$ areas. It is also obvious that 500*500 $m^2$ areas are too large to grasp characteristics of bus stops because the average distance between bus stops is 300 – 500 m. Therefore we need using micromesh to analyze the number of bus users paying attention to a unit of bus stops.

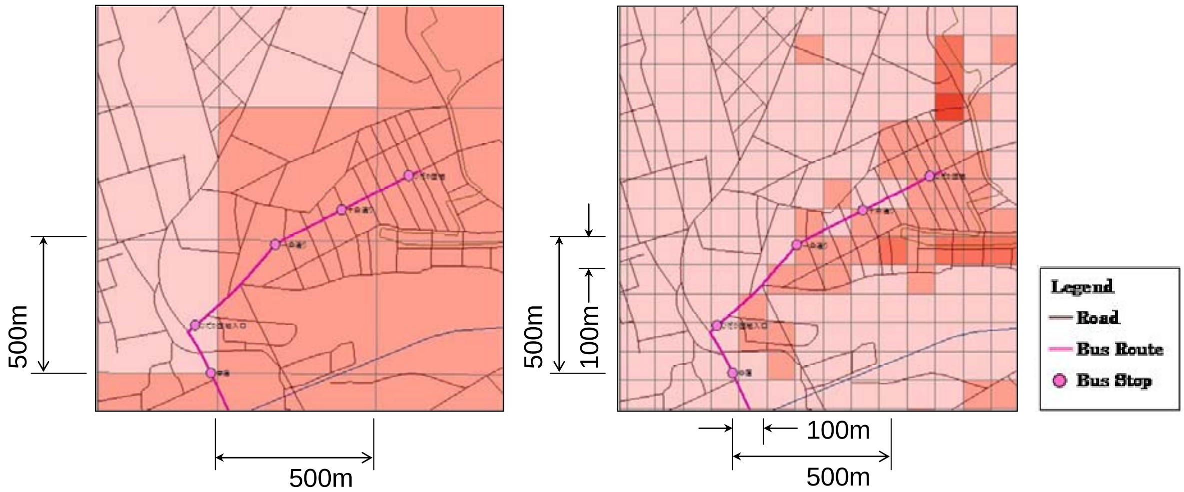


Figure 3 – GIS data used 500*500 $m^2$ (Left) and 100*100 $m^2$ (Right)

Micromesh data is used to be made manually, but it needs huge labor and time or money. The AMS system by Area Market Structure Laboratory Co., Ltd. enables us to make micromesh data automatically. When we make micromesh data, we need to know latitude and longitude of each house for inputting information at position of each house into GIS. We needed to research latitude and longitude of each house on digital map one by one manually until now. It is such a troublesome work that it needs huge cost even if they entrust the work. But the AMS system solved the problem by using digital phone books. The system is followings. First, addresses are researched from telephone number on digital telephone book automatically. Then, addresses are changed into the latitude and longitude coordinates on digital map automatically by geocoder. We can work efficiently by using the AMS system.

## CASE STUDY

This study treats a bus route at Hidaka-city which is located in a suburban area of Tokyo. The city has two large housing estate areas where the number of elderly people is increasing

rapidly. The bus connects these two housing estate areas and three railroad stations (Figure 4). A bus company withdrew because of the management deficit in March 2006, and other bus company, the Eagle Bus Co., Ltd., took over the route afterwards. After 2006, the Eagle Bus carried out questionnaire survey three times in cooperation with Saitama University. As a result, it turned out that the passengers mostly use buses to go to or to come back from each station (Figure 5). In this study it is presupposed that all passengers go to the stations, because the questionnaire survey shows that passengers almost arrive / departure at stations.
Data of population was obtained from a national census 2005 as grid data. Data of bus services; frequency, the required time, fare, and the length of route were collected from the Eagle Bus homepage. We can grasp the collected data visually by using GIS.

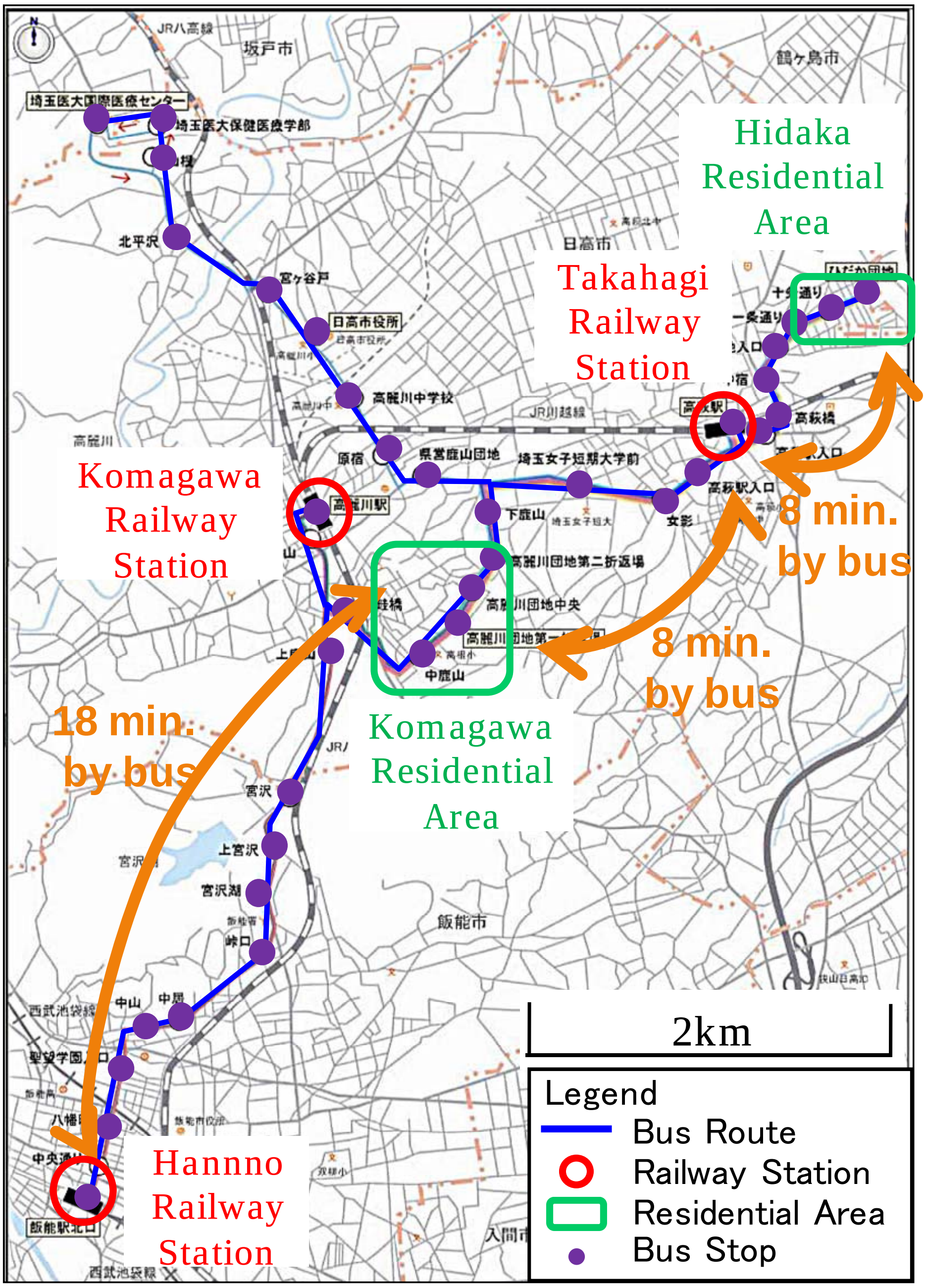

Figure 4 – Map of Hidaka-Hanno route bus

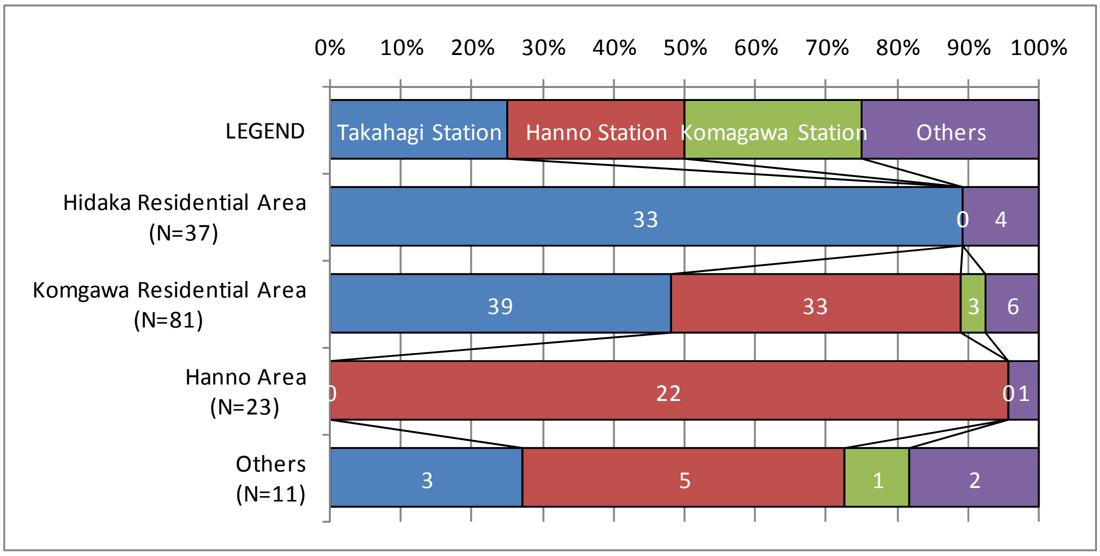


Figure 5 – Destination of bus users at Hidaka-Hanno bus route areas (survey in2006)

## Data on Bus Stop Passenger by Door Sensor

Eagle Bus is trying to accumulate daily passenger data in each bus stop. Japanese bus enterprises don't have to record bus passenger data routinely. When they apply to municipality for the subsidy, they need to submit the data of the number of the passengers in the bus route investigated by one day survey. However, the number of the average daily passenger data made by more than one day survey is not requested for the application. Therefore the attempt of Eagle Bus is exceptional in Japan.
The method is to place sensors at the doors of buses, which count the number of people who are getting on or getting off the bus through the doors at the each bus stop. Basic data such as routes, time, and the number of passengers are important material for planners. One of the characteristics of the system is to grasp the passenger data not with a unit of routes but with a unit of bus stops. They can argue the position of the bus stops which is one of fundamental problems in bus services.

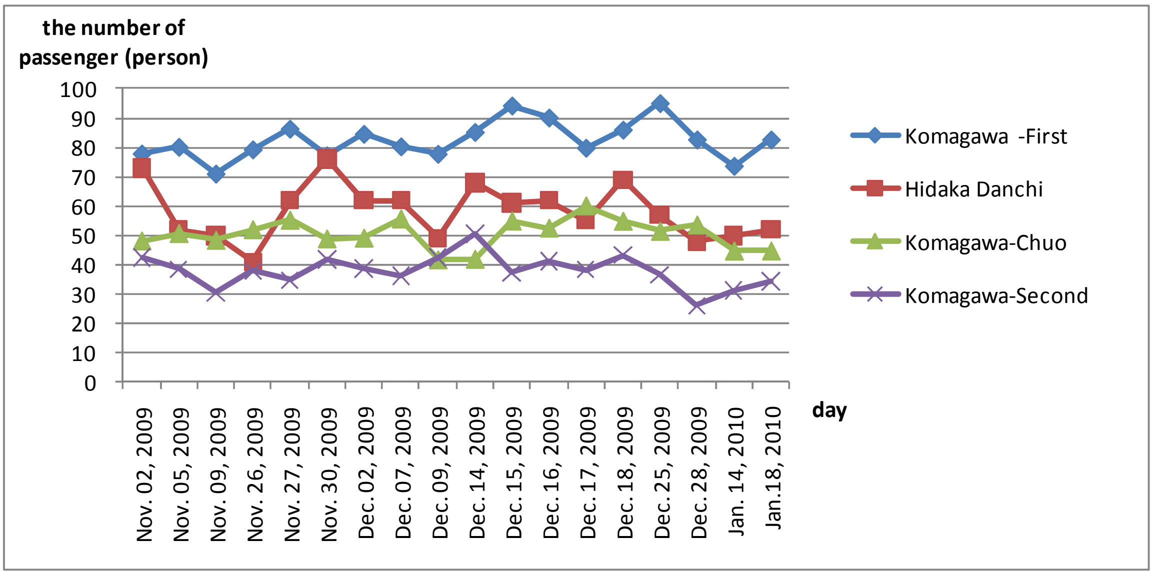


Figure 6 – Fluctuation of passengers per day at the main bus stops

The study uses the number of passenger data by the door sensor as a purpose variable. The Figure 6 shows that the number of the bus passengers changes among days. Because the range of passenger at a total of all bus stops is 133 within the 18 weekdays, it is appropriate that average data made by more than one investigation are used.

## Decision of Bus Stop Territory

The bus stop territory (Figure 7) is considered an area which stands on a circle area with radius 300m except last bus stops. The number of passengers at the last bus stop tends to be more than the normal bus stops, therefore the bus stop territories at the last bus stops are defined area of a circle with radius 500m. The areas are divided by Voronoi Diagram when the areas overlap with the bus stop of the neighbourhoods.

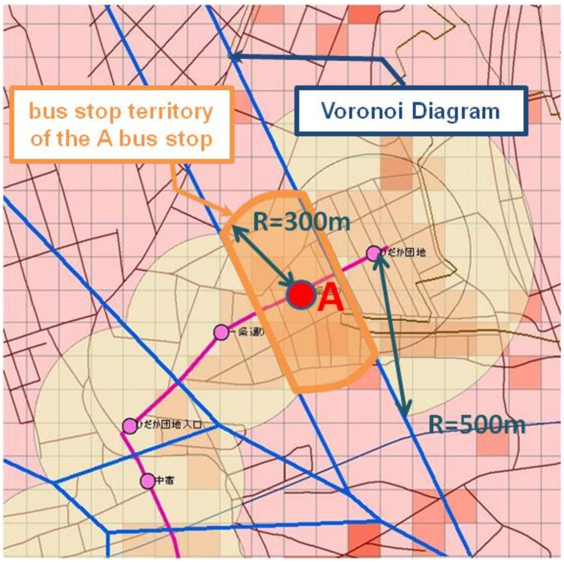


Figure 7–Example of bus stop territory

## Making the BPE Models using Multiple Regression Analysis

We can see the difference of the number of passengers between each hour (Figure 8). The number of passengers getting on from housing estate areas after 17:00 is small. So the study makes two types of the BPE models; one is a model of commute hour (from the first bus to 9:00) and the other is a model of non-commute hour (from 9:00 to 17:00).
Eight variables on Table2 or Table3 are discussed as dependent variables. The dummy variable of walking sphere is defined as follow; if the bus stop is located within a radius of X from the station, the value is 1. Otherwise it is 0. X is 300m, 600m, 900m, and 1200m.
Multiple regression analysis is tried with the following eight cases (Table1). At first the initial best variables are guessed by Case0; stepwise way by SPSS. Next, the most suitable population variable is determined by comparing Case1 - Case3. It is judged by comparison of Case4 - Case7 whether the initial walking sphere variable is the best one, and finally the last BPE model is decided.

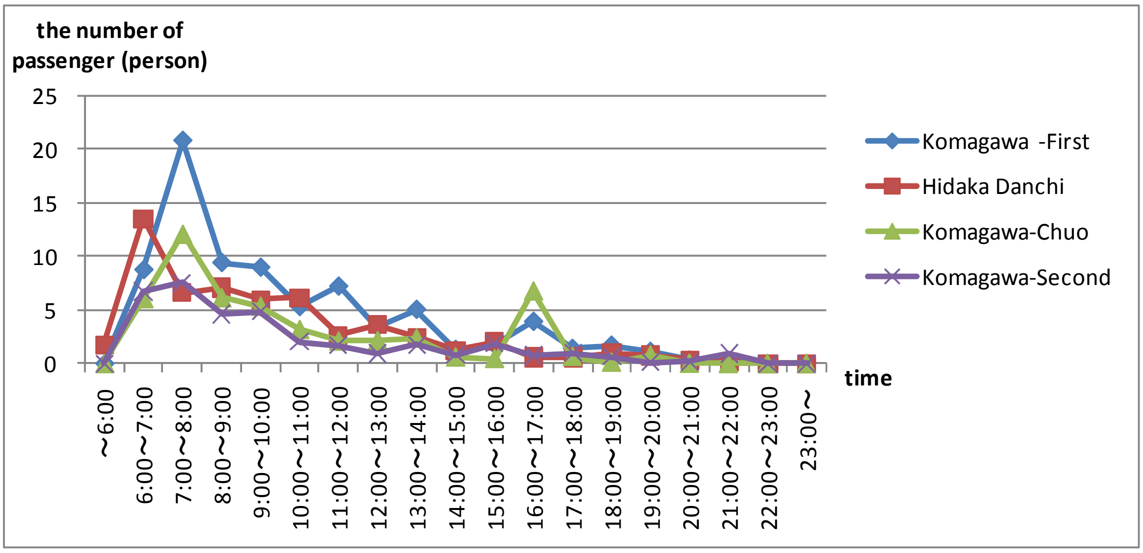


Figure 8– Fluctuation of passengers per hour at the main bus stops

Table 1 – Case of selection of dependent variables

| | |
|---|---|
| Case0 | Selection by stepwise way (SPSS) |
| Case1 | Use productive population as population data |
| Case2 | Use elderly population as population data |
| Case3 | Use employed population as population data |
| Case4 | Use 300m(D) as walking sphere data |
| Case5 | Use 600m(D) as walking sphere data |
| Case6 | Use 900m(D) as walking sphere data |
| Case7 | Use 1200m(D) as walking sphere data |

### *Commute Hour Model*

Table2 is the results of multiple regression analysis concerned commute hour. Compared Case1 - Case3, $R^2$ is biggest and p values of all variables are significant on 1% level. However, sensuously, it is unusable to choose elderly population as a parameter of the commute hour model. Using productive population or employed population is more natural. $R^2$ of Case3 is bigger than that of Case1, so Case3 is adopted.
The final BPE model (commute hour) is;

$$y = -1.821 + 0.019x_3 - 6.324x_6 + 0.372x_8 (R^2 = 0.571)$$

### *Non-Commute Hour Model*

Table3 is the results of multiple regression analysis concerning non-commute hour. Elderly population is desirable to be selected as population variable, because $R^2$ is biggest and p values of all variables are significant on 5% level in Case2 among Case1, Case2, and Case3. Compared Case4 - Case7, $R^2$ of Case5 is slightly bigger than that of Case6. So Case5 is adopted and the 600m (D) is chosen as walking sphere variable.
The final BPE model (non-commute hour) is;

$$y = -3.036 + 0.048x_2 - 5.120x_5 + 0.220x_9 (R^2 = 0.490)$$

Table 2 – Results of multiple regression analysis (commute hour)

| Variable | | | Case0 | Case1 | Case2 | Case3 | Case4 | Case5 | Case6 | Case7 |
|---|---|---|---|---|---|---|---|---|---|---|
| Population | Population of productive age in bus stop territory; x1 | coefficient | 0.014** | 0.014** | | | | | | |
| | | t-value | 3.780 | 3.780 | | | | | | |
| | Population of elderly age in bus stop territory; x2 | coefficient | | | 0.059** | | | | | |
| | | t-value | | | 4.016 | | | | | |
| | Population of the employed in bus stop territory; x3 | coefficient | | | | 0.019** | 0.021** | 0.020** | 0.019** | 0.020** |
| | | t-value | | | | 3.799 | 3.665 | 3.743 | 3.799 | 3.752 |
| a Dummy Variable of Walking Sphere | Distance from bus stop to railway station = 300m (D); x4 | coefficient | | | | | -3.750 | | | |
| | | t-value | | | | | -0.896 | | | |
| | Distance from bus stop to railway station = 600m (D); x5 | coefficient | | | | | | -5.967* | | |
| | | t-value | | | | | | -2.308 | | |
| | Distance from bus stop to railway station = 900m (D); x6 | coefficient | -6.222** | -6.222** | -6.501** | -6.324** | | | -6.324** | |
| | | t-value | -2.995 | -2.995 | -3.216 | -3.057 | | | -3.057 | |
| | Distance from bus stop to railway station = 1200m (D); x7 | coefficient | | | | | | | | -4.431* |
| | | t-value | | | | | | | | -2.130 |
| Level of Bus Service | Frequency per morning time; x8 | coefficient | 0.361* | 0.361* | 0.382** | 0.372* | 0.291 | 0.354* | 0.372* | 0.310* |
| | | t-value | 2.484 | 2.484 | 2.768 | 2.593 | 1.808 | 2.337 | 2.593 | 2.086 |
| | Frequency per daylight time; x9 | coefficient | | | | | | | | |
| | | t-value | | | | | | | | |
| Constant | — | coefficient | -1.721 | -1.721 | -1.995 | -1.821 | -3.529 | -2.889 | -1.821 | -1.564 |
| | Adjusted Decision Coefficient; $R^2$ | | 0.569 | 0.569 | 0.585 | 0.571 | 0.465 | 0.527 | 0.571 | 0.517 |
| | Adjusted Multiple Correlation Coefficient; R | | 0.754 | 0.754 | 0.765 | 0.756 | 0.682 | 0.726 | 0.756 | 0.719 |

*P<0.05, **P<0.01

Table 3 – Results of multiple regression analysis (non-commute hour)

| Variable | | | Case0 | Case1 | Case2 | Case3 | Case4 | Case5 | Case6 | Case7 |
|---|---|---|---|---|---|---|---|---|---|---|
| Population | Population of productive age in bus stop territory; x1 | coefficient | 0.012** | 0.012** | | | | | | |
| | | t-value | 3.308 | 3.308 | | | | | | |
| | Population of elderly age in bus stop territory; x2 | coefficient | | | 0.047** | | 0.051** | 0.048** | 0.047** | 0.047** |
| | | t-value | | | 3.354 | | 3.335 | 3.417 | 3.354 | 3.290 |
| | Population of the employed in bus stop territory; x3 | coefficient | | | | 0.016** | | | | |
| | | t-value | | | | 3.297 | | | | |
| a Dummy Variable of Walking Sphere | Distance from bus stop to railway station = 300m (D); x4 | coefficient | | | | | -1.602 | | | |
| | | t-value | | | | | -0.427 | | | |
| | Distance from bus stop to railway station = 600m (D); x5 | coefficient | | | | | | -5.120* | | |
| | | t-value | | | | | | -2.191 | | |
| | Distance from bus stop to railway station = 900m (D); x6 | coefficient | -3.996 | -3.996 | -4.215* | -4.081* | | | -4.215* | |
| | | t-value | -2.011 | -2.011 | -2.142 | -2.057 | | | -2.142 | |
| | Distance from bus stop to railway station = 1200m (D); x7 | coefficient | | | | | | | | -3.718 |
| | | t-value | | | | | | | | -1.952 |
| Level of Bus Service | Frequency per morning time; x8 | coefficient | | | | | | | | |
| | | t-value | | | | | | | | |
| | Frequency per daylight time; x9 | coefficient | 0.205* | 0.205* | 0.214* | 0.211* | 0.181* | 0.220* | 0.214* | 0.202* |
| | | t-value | 2.488 | 2.488 | 2.649 | 2.584 | 2.118 | 2.713 | 2.649 | 2.508 |
| Constant | — | coefficient | -2.218 | -2.218 | -2.384 | -2.317 | -3.412 | -3.036 | -2.384 | -1.822 |
| | Adjusted Decision Coefficient; $R^2$ | | 0.484 | 0.484 | 0.487 | 0.483 | 0.421 | 0.490 | 0.487 | 0.477 |
| | Adjusted Multiple Correlation Coefficient; R | | 0.696 | 0.696 | 0.698 | 0.695 | 0.649 | 0.700 | 0.698 | 0.691 |

*P<0.05, **P<0.01

## *Comparison of two type models*

Because we intentionally used the employed population data as population data in commute hour model, it is meaningless to argue about the difference of population variables. Only it can be said that the result of Case0 is quite different from what we have expected. We expected that employed population or productive population will be chosen automatically in commute hour model and elderly population will be chosen in non-commute hour. But it was not correct.

Dummy variables of walking sphere might be an intrinsic difference in two models. From the difference of these variables of the two models, it can be said that the users of non-commute hour tend to use bus service even if their houses are close to the station when we compare with the users of commute hour. However, there are little difference between $R^2$ of Case6 and that of Case5 at non-commute hour. That is same as t value. It may be hard to say there are a clear differential between commute hour model and non-commute hour model.

## Application of the BPE Models

The study examines a detour course at Komagawa housing estate areas (Figure 9). The study considers moving the present bus route to the back street of the main way for picking up the inhabitants of the single-family house. Komagawa-Chuo bus stop that is located in the main street will be abolished and new bus stops A, B, and C will be installed in the proposed route. The number of bus stops depends on distance. The bus frequency in each bus stop will not change. The bus frequency at the new bus stops will be same level as an abolished bus stop.

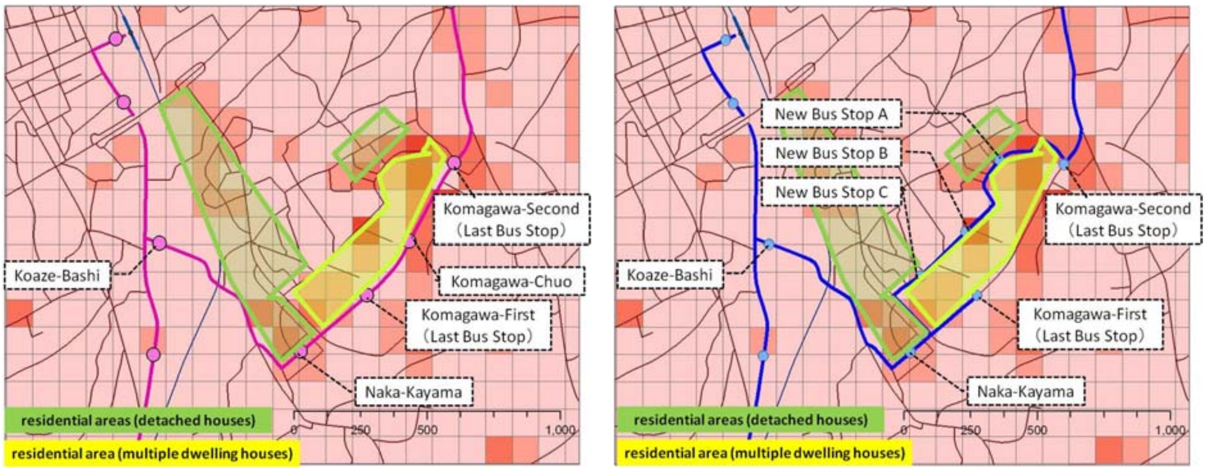

Figure 9– The present route (Left) and the proposed route (Right)

The BPE models are applied for the two routes, and the results are shown in Table4. When each of predicted values is compared with the actual survey value, it indicates that the number of the passengers at Komagawa-First bus stop is underestimated. The Komagawa-First bus stop is the last destination as for the half of the bus which comes from the Takahagi Station. In reality the part of inhabitants near the Naka-Kayama bus stop may use the Komagawa-First bus stop when the Komagawa-First bus stop is the end line. That is to say that the competition with the Naka-Kayama bus stop may influence the result. The territory is just separated by Voronoi Diagram on the GIS therefore the models couldn't consider the real competition.

In the commute hour, there are 16 more passengers in the proposed route than that of the present route. There are 14 more passengers in the proposed route than that of the present route. From the results, it can be said that the number of the passengers will be increased if the bus runs on the proposed detour route.

But the results cannot consider the influence of the extra time which is occurred by the detour. It is estimate that the extra time is less than 3 minute, so the influence is small. But particularly in commute hour, even such little time may be very important for commute people. We have to take it into consideration when we discuss proposed routes.

Table 4 – Comparison of the present route and the proposed route

| | Commuting Hour (the first bus - 9:00) | | | Non-Commuting Hour (9:00-16:00) | | |
|---|---|---|---|---|---|---|
| | Observed Value | Present Route | Proposed Route | Observed Value | Present Route | Proposed Route |
| Komagawa-Second | 19 | 24 | 20 | 15 | 20 | 17 |
| Komagawa-Chuo | 24 | 17 | - | 23 | 14 | - |
| Komagawa-First | 39 | 16 | 16 | 37 | 13 | 13 |
| Naka-Kayama | 7 | 9 | 7 | 6 | 6 | 5 |
| Koaze-Bashi | 0 | 0 | 0 | 9 | 3 | 3 |
| New bus stop A | - | - | 16 | - | - | 13 |
| New bus stop B | - | - | 13 | - | - | 11 |
| New bus stop C | - | - | 10 | - | - | 8 |
| SUM | 89 | 66 | 82 | 90 | 56 | 70 |

# CONCLUSION

This study have made the BPE model using population and bus service level to examine which bus stop locations can increase the number of passengers as one of approach to improve bus businesses in a suburb housing estate area. It is dangerous to swallow a mathematically calculated result and plan a bus route. However, the result is good for smooth planning of bus services when it is used like one of information. This study reveals that the BPE model can help an enterprise to examine the improvement of bus route concretely considered as the one technique.

But the decision coefficient $R^2$ of the model has to improve more in future. For example, the bus stop territory is established on only an area of a circle, but the model may have higher adequateness if a distance from someone's house to a bus stop is considered in it. In addition the models must be improved to be a flexible model for application of various areas.

# REFERENCE

Satoshi FUJII (2003) Travel Demand Forecasting using Behavioural Intention Method an Application Study for Latent Demand Forecasting of New Bus Route, Infrastructure Planning Review by Japan Society of Civil Engineers, 20(3), 571-578

Itsuki YOSHIDA, Tetsuo AKIYAMA (2004) Studies on Planning of Bus Service Standards Considering Potential Travel Needs in the Urban Areas, Journal of the City Planning Institute of Japan, No.39-3, 81

Keita SUGIO, Tomohiko ISOBE, and Denshi TAKEUCHI (2001) Development of Supporting System for Bus Network Planning Applying GIS –Bus Route Assessment based on Potential Demand-, Infrastructure Planning Review by Japan Society of Civil Engineers, 18, 617-626