Microarea Selection Method for Broadband Infrastructure Installation Based on Service Diffusion Process

Motoi Iwashita, Akiya Inoue
Dept. of Management Information Science
Chiba Institute of Technology
Chiba, Japan
email:{iwashita.motoi, akiya.inoue}@it-chiba.ac.jp

Takeshi Kurosawa
Dept. of Mathematical Information Science
Tokyo University of Science
Tokyo, Japan
email: tkuro@rs.kagu.tus.ac.jp

Ken Nishimatu
Network Technology Laboratories
NTT
Tokyo, Japan
email: nishimatsu.ken@lab.ntt.co.jp

***Abstract*—Wired/wireless information communication networks have been expanded to meet the demand of broadband services as part of information and communication technology (ICT) infrastructure. As the installation and expansion of ICT infrastructure requires a large amount of time and money, the decision on how to select the installation area is a key issue. Low-usage facilities can cause problems for businesses in terms of investment efficiency. Moreover, it takes time to select areas because of the need to estimate the potential demand and to manage the installation of the infrastructure for thousands of municipal areas across a nation. In this paper, we propose an efficient microarea selection method for use during the life cycle of broadband services, i.e., from early to late stage. This method is developed considering consumer segmentation, the broadband service diffusion model by consumer behaviour, and area characteristics based on employee fluidity. The proposed method is evaluated on worldwide interoperability for microwave access (WiMAX) and its applicability is ascertained on the basis of the infrastructure's area penetration rate and area characteristics.**

Keywords-broadband services; infrastructure installation; data mining; area marketing; area characteristics; demand forecast; decision support system; algorithm.

I. INTRODUCTION

Broadband access services can be rapidly deployed by asymmetric digital subscriber line (ADSL) penetration. This has enabled consumer-generated media (CGM) such as social networking service (SNS) and YouTube to be widely used on a broadband access infrastructure. These multimedia services have dramatically changed the modern lifestyle and made it possible for individuals to obtain and share information with ease. The research of broadband infrastructure installation has been done [1] for providing these services. Some local governments are trying to utilize ICT infrastructure for healthcare and nursing among other applications in their respective regions [2]. Many companies have also introduced ICT elements such as mobile gadgets for sales, maintenance, operation, and production to improve the efficiency of corporate functions.

Wired broadband access infrastructure has propagated rapidly first by the use of ADSL and then optical fibres in the fibre-to-the-home (FTTH) infrastructure. FTTH is an ultra-high-speed broadband access infrastructure, which is being provided in Japan since 2002. Such an ICT infrastructure provides a variety of technologies and benefits for corporate activity. Although the coverage rate of FTTH as a percentage of the total number of households nationwide in 2014 was approximately 98% [3], the customer rate (percentage of customers using FTTH) in the coverage area was merely 43%. With respect to wireless broadband access, long-term evolution (LTE) for high-speed wireless access is being provided in Japan since December 2010. The coverage rate of LTE was more than 90%, while the customer rate was approximately 42% in 2014 [4]. The other high-speed wireless access worldwide interoperability for microwave access (WiMAX), which is being provided since 2009, had a coverage rate of greater than 90%, and 7 million users in 2014.

Because installation of infrastructure is capital intensive, business profitability is significantly impacted if the facility usage is low. It is difficult to identify low-usage areas when we focus on the average data. This demonstrates the importance of considering not only macro areas but also micro areas when installing the infrastructure. Therefore, strategic and economic considerations are necessary for the installation of ICT infrastructure, such as broadband and wireless access facilities. Such installations greatly depend on the potential demand in different microareas. Further, the existence of more than a thousand microareas, such as municipal areas, necessitates using an efficient estimation method.

The goal of providing an area with ICT services is to determine the investment order of microareas, where the ICT infrastructure is installed several months prior to the installation. ICT infrastructure installation per area is more effective and less expensive than on-demand installation in which a facility is installed on a case-to-case basis. Furthermore, a method to efficiently select a microarea will have a positive impact on the operations and financial efficiency of an enterprise.

In this paper, we propose a microarea selection method that is simpler to use than trade analysis [5]. Our target is ICT infrastructure installation during the life cycle, which covers different stages of ICT infrastructure installation for broadband services. The proposed method is based on consumer segmentation based on different consumer behaviours in the broadband service propagation model and area characteristics based on employee fluidity. We verified the proposed model with the penetration of WiMAX services. The remainder of the paper is organized as follows: Section II introduces related works. Section III describes the trend of WiMAX demand in Japan. Section IV discusses the hypothesis of service diffusion and its model. Section V describes the proposed method for the selection of microareas. Section VI presents the simulation results and evaluates the method. Section VII clarifies the applications of the proposed method based on area characteristics. Section VIII concludes this paper.

II. RELATED WORKS

The determination of the target area for marketing, such as areas to focus on for sales activities and areas to install the facility in, is based on trade analysis. Trade analysis is a well-known method for the investigation of geographical areas for business deals and involves demographic data and field surveys. For instance, the setting up of a convenience store is decided on the basis of demographic data and field surveys. Geographic information system (GIS) [6] is an effective tool for area-related decision making. An empirical study using GIS for trade analysis has been previously [7] reported and many companies uses this method with map data. These approaches are effective for deciding whether a store should be set up in a given area.

Application of these methods to ICT infrastructure installation requires spending a large amount of time on selecting areas. This is because installation in only one target area has little effect from the viewpoint of network externality [8], which makes it more convenient to have more users, if it is carried out in one area rather than in several areas simultaneously and nationwide. The idea of modelling spatial data that represent a geographical location has been proposed earlier [9], [10]. This approach is now being practically used to understand geographical features.

To select the areas, we need to first consider potential demand. Previous research [11]-[13] has focused on macro demand forecast to provide facility installation principles and to not select installation areas. There has been one study on microarea forecasting [14]. It describes only the guidelines for microarea forecasting by using multiple regression analysis.

To proceed with microarea marketing, we focused on the expansion of ICT infrastructure in accordance with service reputation in areas where ICT infrastructure has been provided as a trial. In such cases, individuals tend to exhibit a “go type behaviour” where individuals tend to go to issues/people when forming preferences [15], [16]. Therefore, who pushes service forward is an important question. An innovative early adopter in the technology lifecycle is characterized as an information source affecting acquaintances from the viewpoint of innovative diffusion [17]. The innovative early adopter is generally reckoned as a person who gets stimulated with many contacts through his/her mobility. It is difficult to characterise each person in an area from the viewpoint of the technology lifecycle (e.g., who is an early adopter). Recently, a city planning study that uses mobile phones (life log data) to obtain mobility data by area has been initiated [18]. The researchers expect to analyse human behaviour by utilizing the life log data, and if it yields successful results, it can be applied to various fields. However, currently, there is no such information available.

Concerning the diffusion of the broadband infrastructure facilities at the moment, the framework of microarea marketing is based on commuting flows in terms of considering human behaviour [19]. Such a flow-based microarea selection method has been developed and been compared to the population-based method, which is a simple application of the population order in some regions as the case study [20]. Although these studies show the efficiency of selecting microareas, the results obtained are not stable; i.e., setting up the conditions in advance is difficult. The condition for the application of flow-based microarea selection method is decided empirically and analytically [21]. However, the application of this method is limited to only the early stage of providing broadband services, which means that the area penetration is low. Therefore, a mixed algorithm that consists of both the flow-based and population-based microarea selection methods is proposed that is applicable throughout the life cycle (from early to late stage) of broadband services [1]. However, the obtained results are not stable, i.e., the cause that the mixed algorithm gives the optimal order of microareas depends on the area itself. Therefore, conditions considering not only consumer behaviour but also area characteristics are needed.

III. WIMAX SERVICES AND TREND IN JAPAN

WiMAX is a wireless broadband access service that has been in place in Japan since 2009. It is defined in IEEE 802.16-2004 as an international standard with a maximum transmission distance of 50 km and maximum transmission rate of 70 Mbps. WiMAX network is explained in Fig. 1. Mobile gadgets such as smartphones and ultra-mobile PCs can access the nearest base station by air in each region. The base station transmits signals to the providers’ server through the carrier’s network. The user can then make use of wireless high-speed internet access services. Table I shows in chronological order the events that are related to WiMAX service diffusion. In seven months after WiMAX started to be commercially available, 5,000 base stations were installed nationwide. The pace of base station installation slowed until April 2012 when providers achieved 20,000 base station installations. The number of customers grew steeply from 2009 to 2013, and it took only five months for the growth from 2 to 3 million customers. In October 2013, the new WiMAX service, which could achieve a much higher high-speed transmission rate, was introduced. It was simultaneously provided by mobile virtual network operators (MVNOs). An MVNO is defined as a network service operator that does not have a network facility itself, but rather borrows the facility from a real network operator. Therefore, new MVNOs provide low-priced services and value-added services, such as character brand gadgets and rich video content. Customers have many options for network operators through SIM-free terminals. As for the base station installation for new WiMAX service, the installation pace is shorter than that of WiMAX.

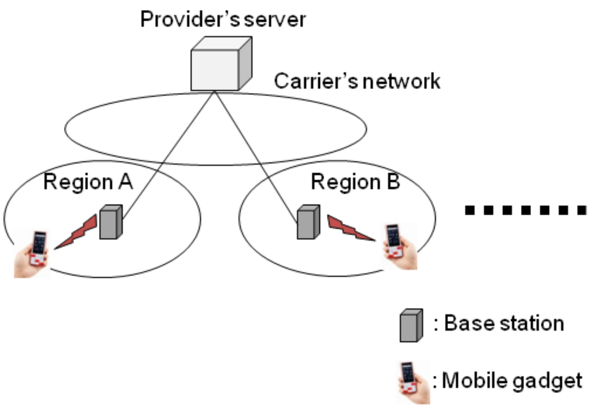

Figure 1. WiMAX network.

TABLE I. EVENTS IN CHRONOLOGICAL ORDER

Date	Event
Feb/2009	WiMAX trial (cost-free) service started
Jul/2009	WiMAX commercial service started
Jan/2010	5,000 base stations installed
Aug/2010	10,000 base stations installed
May/2011	15,000 base stations installed
Jun/2011	1,000,000 customers
Feb/2012	2,000,000 customers
Apr/2012	20,000 base stations installed
Jul/2012	3,000,000 customers
Feb/2013	4,000,000 customers
Oct/2013	New WiMAX service (ultra-high speed) started/ MVNO started
Feb/2015	20,000 base stations installed for new WiMAX service

The total demand for WiMAX has been increasing, with about 20 billion customers in 2015 as shown in Fig. 2. The demand increased at a consistent rate until 2014, and then increased sharply from 2014 to 2015. This is because MVNOs were introduced in mid-2013. MVNO*1 represents the demand excluding real mobile network operators, while MVNO*2 represents the demand including the results from real mobile network operators as MVNOs. These results show that the effect of MVNO*1 on the total demand is smaller, while the total demand is almost the same with MVNO*2. Considering these results, customers who have already been WiMAX users do not change their service to MVNO, this is because the total demand is almost the same as that of MVNO*2. Therefore, newer customers tend to choose their MVNO in terms of pricing compared to 2013.

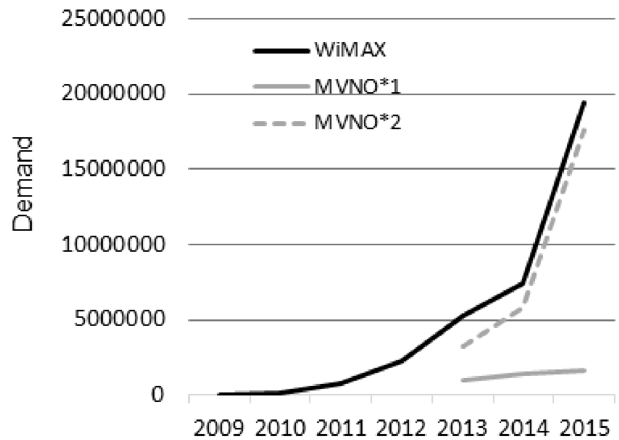

Figure 2. Demand for WiMAX in Japan.

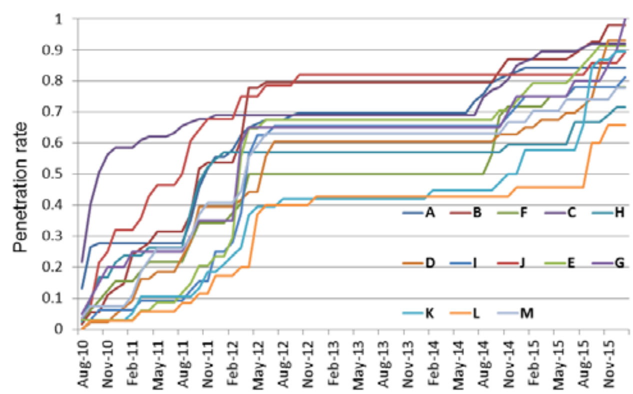

Figure 3. Trend for number of microareas for WiMAX.

Fig. 3 shows the trend of penetration of microareas for WiMAX from August 2010 to November 2015. A microarea is defined as the municipal area in this study. The vertical axis is the penetration rate of WiMAX services, which means the ratio of the number of microareas for a WiMAX facility installed against the total number of microareas in the prefecture. Thirteen prefectures were considered and defined as A to M. The graph shows that many microareas were selected and that the facilities were installed from August 2010 to April 2012. This time interval corresponds to the early stage of WiMAX diffusion. The demand grew inside the areas from 2012 to 2014 since the number of installed microareas did not change. In late 2014, the number of installed microareas began to increase. This situation corresponds to the introduction of MVNO and this means that the customers who are interested in the price of services are major factors for demand increase.

IV. HYPOTHESIS OF SERVICE DIFFUSION AND ITS MODEL

In this section, the assumption of how broadband services are propagated is explained. Many ICT services are already being provided. Therefore, we investigated the features of the propagation of such services. For example, analysis of the mechanism of word of mouth has been studied [22]-[24]. Let us first consider the terminal equipment. The diffusion of audio-visual (AV) and digital equipment depends on the users' experiences through their use and the sharing of their experience with their friends and families by word of mouth. Since many people refer to the site of collecting word-of-mouth information, consumers tend to be affected by a person who has similar preferences.

Broadband services based on the ICT infrastructure become more widespread owing to the sharing of information among users of the AV and digital equipment such as smartphones and tablet PCs. CGM depends significantly on network externality; for example, the ratio of users who choose internet video sites on the basis of recommendations of friends/acquaintances was found to be about 38% [25]. Network externality is defined as the phenomenon in which the benefit of the customers is greater with a greater increase in the number of customers, particularly in terms of networked services.

If an individual has to pay for a new software application or a service upgrade, he/she tends to decide on the basis of face-to-face information from friends/acquaintances [26]. Since the wired/wireless broadband access infrastructure is not a free service, we expect customers to respond to it in the same way they respond to software and service upgrade purchases.

There are five types of consumers, namely the innovator, early adopter, early majority, late majority, and the laggard. An early adopter is a trend-conscious person who collects information, makes decisions by himself/herself, and plays an important role for service diffusion. It is widely known that an early adopter has a considerable influence on general consumers as an opinion leader. This implies that he/she sends interesting information to his/her acquaintances. We made a hypothesis of service diffusion in early stage on the basis of personal behaviour [21]. As face-to-face communication with friends/acquaintances is the key of broadband service diffusion, it is necessary to introduce the concept of innovation diffusion [17]. If an individual (especially early majority) has many contacts with early adopters, the possibility that he/she will demand the service is high [27], as shown in Fig. 4 (a).

Early majority is commonly a person who is easily affected by not only the early adopter but also the affected majority [28]. This implies that each early majority changes his/her mind based on the advice from friend/family/acquaintance. Therefore, face-to-face communication in neighbourhood influences the increase in early majority, as shown in Fig. 4 (b).

The late majority tends to be sceptical of the services, price-oriented and follow the maturity of the early majority after the service has sufficiently penetrated the field. Therefore, the late majority is not affected by contact with the early majority through face-to-face communication, but rather by resonance [29] with early majority, as show in Fig. 4 (c).

Next, we explain the mechanism of potential demand diffusion physically, i.e., what kind of microareas have quick service diffusion. Basically, there are two kinds of service propagation conditions. One is population in a microarea. Higher the population in an area, higher the possibility of increased contacts (face-to-face communication). The other is the in-/out-flow such as human fluidity of microarea. Higher the flows in an area, higher is the possibility of contacts.

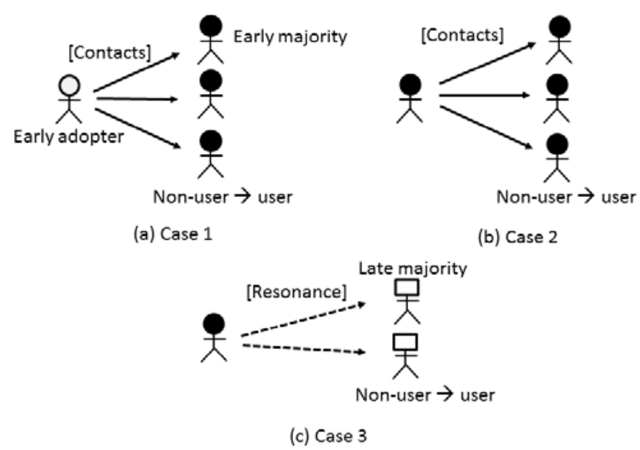

Figure 4. Relationship between service diffusion and customer segmentation.

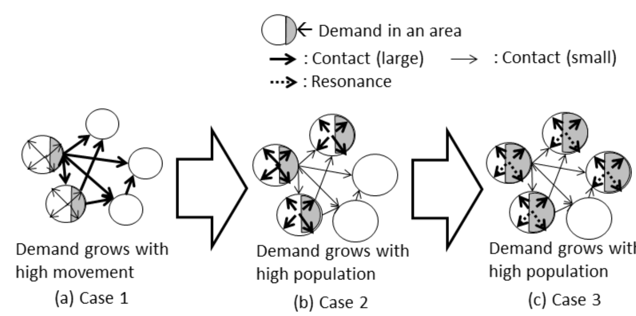

Figure 5. Mechanism of potential demand diffusion.

As face-to-face communication is important in the diffusion of wired/wireless broadband services, the mobility of an early adopter may influence broadband service diffusion through contacts with many individuals in Case 1 (see Fig. 5). Therefore, our hypothesis of Case 1 is that demand grows faster in case of high movement among microareas compared to the case of high population in a microarea, as shown in Fig. 5 (a).

Although face-to-face communication is still effective in Case 2, the contacts among early majority are dominant. Therefore, the service is diffused by friends/acquaintances/families nearby location. Our hypothesis of Case 2 is that the demand grows with high population as shown in Fig. 5 (b). If the differences of population among microareas are small, high movement still results in service diffusion.

In Case 3, resonance leads to the increase in late majority. This means that diffusion does not depend on face-to-face communication. The deeds of early majority decide the deeds of late majority. Therefore, the effect of population in microareas becomes great when compared to that of movement among microareas as shown in Fig. 5 (c). If the differences of population among microareas are small, the same situation occurs in Case 2.

Next, the movement between microareas is defined. Since the diffusion of ICT infrastructure strongly depends on the application (SNS, Net-Game, etc.), it would be desirable to identify commuting flow on the basis of attributes (employee, student, etc.). In addition, if we could access life log data, we could perform an even more detailed analysis. For the frequency of collecting commuting flow information, we assume that an average commuting flow per day or yearly can be used. This is because the interval of ICT infrastructure installation is not frequent (no real-time installation). Therefore, movement between microareas as shown in Fig. 6 (a) is assumed to be modelled applying commuting flow between microareas (Fig. 6 (b)).

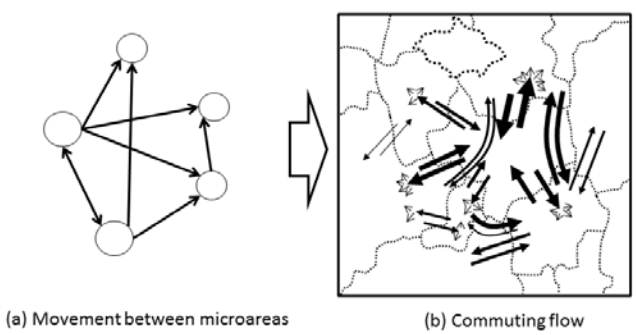

Figure 6. Definition of movement between microareas.

Now we discuss the relationship between our three cases and WiMAX service diffusion. The early stage of service diffusion corresponds to the time interval between 2010 and 2012 according to the results in Figs. 2 and 3. The penetration rate grew with a steep slope and the number of microareas increased sharply. In this stage, the early majority was affected by early adopters as Case 1. The middle stage of service diffusion corresponds to the time interval between 2012 and 2014. In this stage, there was no increase in the number of installed microareas. This means that the early majority who were affected by early adopters contacted others in the early majority in the same microarea. Thus, the demand increased inside a microarea as seen in Case 2. The penetration rate increased in all areas during the late stage of service diffusion after 2014. The new type of services provided by MVNO started in this stage, where there was potential demand from the late majority. The demand grows inside a microarea and this results in addition of new microareas, therefore, the number of micoareas increases. This stage corresponds to Case 3.

V. Microarea selection algorithms

Let us introduce three algorithms to select microareas in this section.

A. Flow-based algorithm

To select microareas on the basis of the movement, we created a table of the inflows and outflows to and from microareas based on commuting flows among microareas, as shown in the table presented in Fig. 7.

Table of in- and outflows among microareas

Out \ In	Area 1	Area 2	Area 3	…
Area 1	f_{11}	f_{12}	f_{13}	
Area 2	f_{21}	f_{22}	f_{23}	
Area 3	f_{31}	f_{32}	f_{33}	
⋮				

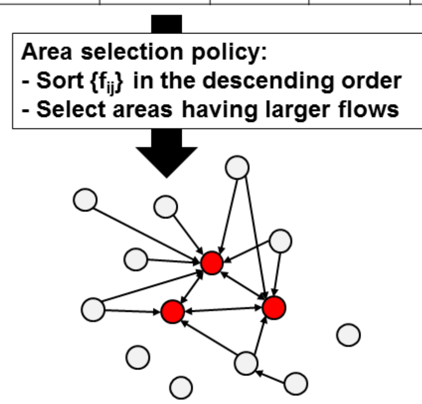

Figure 7. Concept of flow-based algorithm.

Let these flows be sorted in the descending order to estimate the area selection efficiently. Furthermore, we built a graph model in which a microarea is denoted as a node and a commuting flow among the microareas is denoted as a link. Each link has an arrow to indicate the direction of the commuting flow. A link was added among the specified areas if the in/outflow was greater than or equal to the given threshold, α, penetration rate. In other words, we selected links that contained a high average number of commuters as shown in the lower part of Fig. 7.

Thus, a microarea selection method using a flow-based algorithm was constructed according to the above procedure, as shown in Fig. 8. Its definition and notation are as follows:

Let $G = (N, L)$ be a graph, where N denotes a finite set of nodes i ($i \in N$), and L represents a set of links l_{ij} ($l_{ij} \in L$, i, $j \in N$).

Let f be a function such that $f(l_{ij}) = z_{ij}$ in N_0 (non-negative integer) for $l_{ij} \in L$.

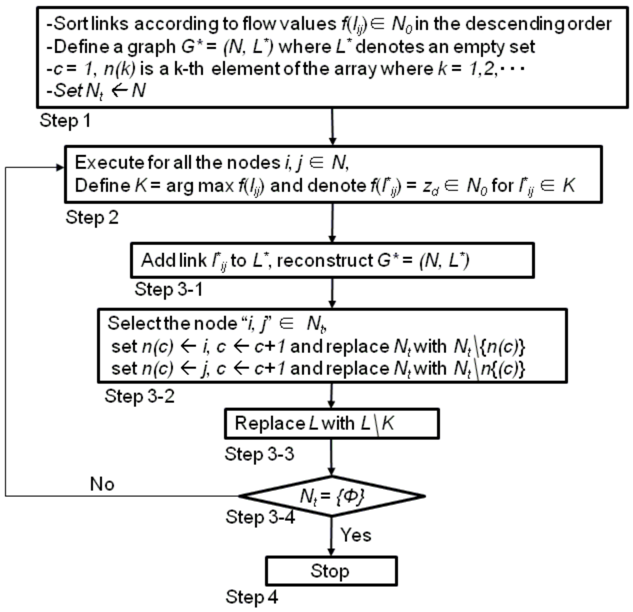

Figure 8. Flow-based algorithm.

B. Population-based algorithm

To select microareas on the basis of population, we created a table of the population of each microarea, as shown in the table presented in Fig. 9.

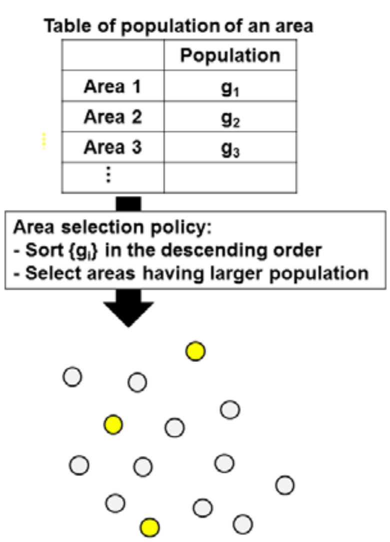

Figure 9. Concept of population-based algorithm.

Let these values be sorted in the descending order to select the node with larger values as shown in the lower part of Fig. 9.

Thus, microarea selection method using population-based algorithm was constructed according to the below procedure, as shown in Fig. 10. Its definition and notation are as follows:

Let $G = (N, L)$ be a graph, where N denotes a finite set of nodes i ($i \in N$).

Let g be a function such that $g(n_i) = u_i$ in N_0 (non-negative integer) for $n_i \in N$.

Therefore, g gives population in each microarea.

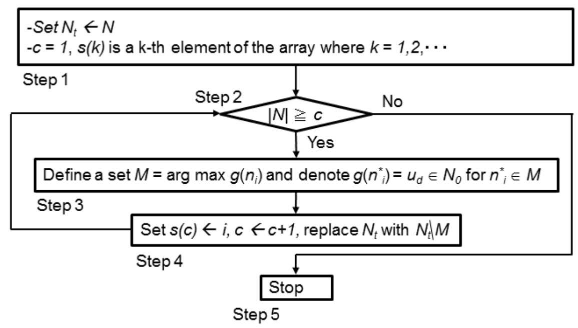

Figure 10. Population-based algorithm.

C. Mixed algorithm

The population-based algorithm is defined as the method that selects areas by the application of the population order; therefore, areas with a large population tend to be selected. The flow-based algorithm is defined as a method that selects areas on the basis of the inflows and outflows among areas. According to the mechanism described in the previous subsections, the mixed algorithm is constructed so that the flow-based algorithm is used in the early stage while the population-based algorithm is used in the middle and late stages as shown in Fig. 11.

Let $G = (N, L)$ be a graph, where N denotes a finite set of nodes ($i \in N$) and L represents a set of links ($l_{ij} \in L$).

There exists a function f: $L \rightarrow N_0$ (non-negative integer) such that $f(l_{ij}) = z_{ij}$ for any $l_{ij} \in L$, and a function g: $N \rightarrow N_0$ (non-negative integer) such that $g(n_i) = u_i$ for $n_i \in N$.

The procedure to construct the mixed algorithm is as follows;

- Step 1: Let c be a counter with an initial value '1', and $n(k)$ be the k-th element of the array where k = 1, 2,… Let "p" denote the number of selected areas by WiMAX evolution under the penetration rate (α%), Set $N_t \leftarrow N$.
  Sort links according to flow values $f(l_{ij})$ in the descending order.
  Sort nodes according to population values $g(n_i)$ in the descending order.

- Step 2: While $p \geq c$, then the following steps are performed (flow-based algorithm):
    - Step 2-1: Define a set K ($\subset L$) = arg max $f(l_{ij})$, and denote $f(l^*_{ij}) = z_d \in N_0$ for $l^*_{ij} \in K$.
    - Step 2-2: Select nodes 'i, j' $\in N_t$, then set $n(c) \leftarrow i$, $c \leftarrow c+1$ and replace N_t with $N_t \backslash \{n(c)\}$, and set $n(c) \leftarrow j$, $c \leftarrow c+1$ and replace N_t with $N_t \backslash \{n(c)\}$.
    - Step 2-3: Replace L with $L \backslash K$.
- Step 3: $c \leftarrow p + 1$.
- Step 4: While $|N| \geq c$, then the following steps are performed (population-based algorithm):
    - Step 4-1: Define a set M ($\subset N_t$) = arg max $g(n_i)$, and denote $g(n^*_i) = u_d \in N_0$ for $n^*_i \in M$.
    - Step 4-2: Select node 'i', and set $n(c) \leftarrow i$, $c \leftarrow c+1$, and replace N_t with $N_t \backslash M$.

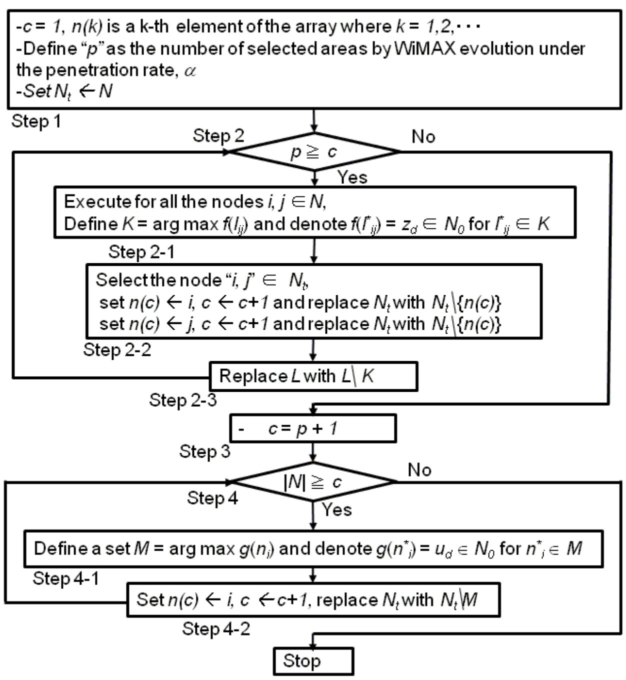

Figure 11. Mixed algorithm.

VI. EVALUATION OF ALGORITHMS

Firstly, microareas are classified according to employee fluidity, then three algorithms are compared. Moreover, the relationship between algorithms and area characteristics are described.

A. Classification of prefectures

As a microarea usually belongs to a prefecture, it is better to target microareas in a specified prefecture. Forty-one prefectures are considered to be representative of all the prefectures in Japan in terms of the population size. However, the following two types of prefectures are considered to be exceptions and are excluded from this study:

- Prefectures with a very large population, such as Tokyo, Osaka, Kyoto, and Kanagawa, which have high fluidity and a high diffusion speed in their own region
- Prefectures such as Hokkaido and Okinawa, which are islands with low fluidity as compared to the other prefectures.

As employee fluidity (as commuting flows) depends on the area selected especially in the early stage [20], the attributes of prefectures are considered to be 'number of employees in own microarea', 'inflow of employees among microareas', and 'outflow of employees among microareas'. We classified the prefectures into the following five categories in terms of employee fluidity according to the correspondence analysis shown in Fig. 12. The vertical axis represents the occurrence ratio of the in- or outflow among the microareas in the prefecture. The horizontal axis represents the staying ratio of employees in their own microareas.

- Group 1: areas with large inflow; Aichi (A)
- Group 2: areas with large in- and outflows; Ibaraki (B) and Shiga (F)
- Group 3: areas with large outflow; Saitama (C) and Nara (H)
- Group 4: areas with staying in own area as little movement; Niigata (D), Okayama (I) and Hiroshima (J)
- Group 5: balanced areas of average in- and out-flow; Kagawa (E), Ishikawa (G), Yamagata (K), Tokushima (L) and Fukui (M)

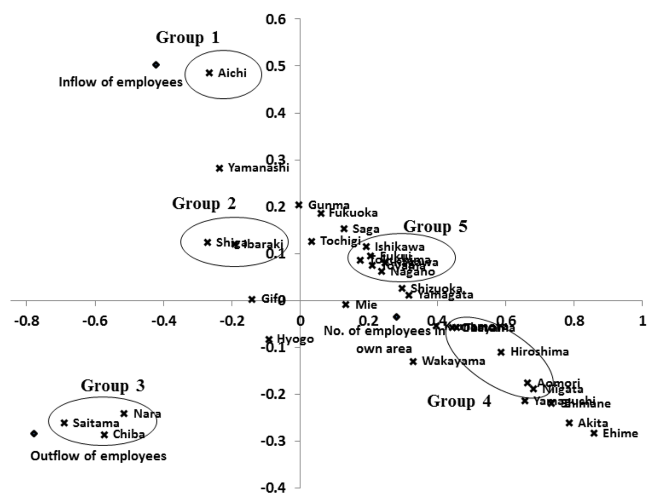

Figure 12. Correspondence analysis of employee fluidity.

B. Comparison with algorithms

In this section, we compare and evaluate the population-based, flow-based and mixed algorithms. To determine the difference in results depending on region, thirteen prefectures were considered among five groups (A to M).

Since infrastructure installation takes a large amount of time, a yearly plan of the area installation order is necessary. The penetration rate is calculated as the ratio of the number of areas, where WiMAX has been introduced by the provider (WiMAX evolution), to the number for all areas in the given prefecture. We take two conditions as penetration rate: 40% and 80% corresponding to low and high stage, respectively.

Table II shows the concordance ratio (CR) comparison among flow-based and population-based algorithms when the penetration rate is 40%. The CR of the number of selected areas between WiMAX evolution and each algorithm is defined by the following equation.

TABLE II. COMPARATIVE RESULTS AT LOW PENETRATION

Prefecture (Area no.)	Group	Population-based algorithm	Flow-based algorithm
A (83)	1	0.62	0.72
B (54)	2	0.6	0.65
F (32)	2	0.77	0.85
C (87)	3	0.69	0.69
H (42)	3	0.81	0.81
D (43)	4	0.53	0.71
I (32)	4	0.58	0.67
J (28)	4	0.71	0.71
E (34)	5	0.79	0.79
G (20)	5	0.57	0.86
K (38)	5	0.73	0.87
L (35)	5	0.79	0.79
M (27)	5	0.73	0.82

$$\text{CR} = (\text{Number of selected areas matching WiMAX evolution areas} / \text{Number of WiMAX evolution areas}). \quad (1)$$

The underlined values (in Table II) indicate the highest CR at each prefecture. Although the CR by population-based algorithm is sometimes the highest, the CR by flow-based algorithm is always the highest for all Groups. Therefore, flow-based algorithm is suitable in the early stage.

Table III shows the concordance ratio comparison among the three algorithms. Table III results are for the penetration rate of 80% in each prefecture. The mixed algorithm works such that flow-based algorithm is used when the penetration rate reaches 40%, population-based algorithm is used beyond 40%.

Although the CR by flow-based algorithm is sometimes the highest, the CR by population-based algorithm is always the highest for Groups 1 to 4. Employee fluidity explains well, the behaviour in the early stage by applying the flow-based algorithm, while the resonant effect well explains the demand increase by population in the late stage by applying the population-based algorithm. Therefore, the mixed algorithm is for use in accordance with the penetration rate for Groups 1, 2, 3 and 4 during the life cycle of the services. However, the CR by flow-based algorithm is always superior to that by population-based algorithm for Group 5. It is better to use the flow-based algorithm for the whole penetration rate.

TABLE III. COMPARATIVE RESULTS AT HIGH PENETRATION

Prefecture (Area no.)	Group	Population -based algorithm	Flow-based algorithm	Mixed algorithm
A (83)	1	0.94	0.92	0.94
B (54)	2	0.79	0.79	0.79
F (32)	2	0.80	0.80	0.80
C (87)	3	0.93	0.91	0.93
H (42)	3	0.97	0.93	0.97
D (43)	4	0.82	0.82	0.82
I (32)	4	0.92	0.88	0.92
J (28)	4	0.91	0.86	0.91
E (34)	5	0.81	0.89	0.81
G (20)	5	0.81	0.81	0.81
K (38)	5	0.81	0.87	0.81
L (35)	5	0.87	0.91	0.87
M (27)	5	0.86	0.90	0.86

C. *Consideration for microarea characteristics*

In this subsection, we consider the differences between Group 5 and the other groups to apply the proposed algorithm. The differences of population between areas are focused and analysed. Figs. 13, 14, 15, 16, and 17 show the relationship between the population in a microarea, and ranking of microareas for prefectures. The target microareas excluded the microareas that were selected by WiMAX evolution when the penetration rate was lower than α (α = 40% in this study) in each prefecture.

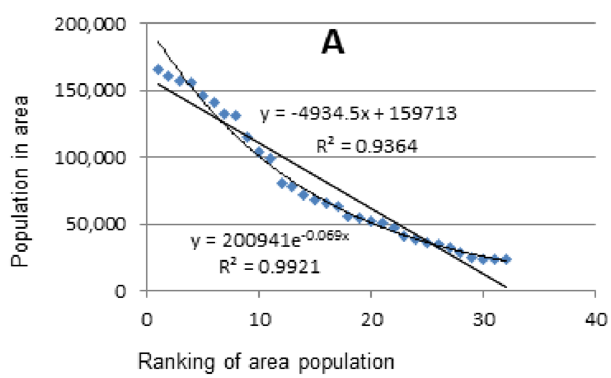

Figure 13. Relationship between population in the area and its ranking (Group 1).

Three approximation curves and their R-square (R^2) values are calculated as exponential, logarithmic, and linear regressions. The result showed that the approximation by exponential regression fits into the scatter diagram at the highest R^2 (R^2 = 0.9921) in A (Group 1), while the approximation by linear regression has R^2 = 0.9364 as shown in Fig. 13. There are many microareas and their population is considered to be linearly decreasing according to the regression line, since R^2 difference of two regressions is small within the given microareas. Therefore, the effect by population difference is large because of linearity. Thus, the population-based algorithm works well in the late stage.

Next, the scatter diagrams in C and H (Group 3) are shown in Fig. 14. The results showed that the approximation by exponential regression also fits into them at the highest R^2 = 0.9834, 0.9493 in C and H, respectively. Their population is considered to be linearly decreasing according to the regression line, since the R^2 difference of two regressions is small with the given microareas. Therefore, the effect by population difference is large because of linearity. Thus, the population-based algorithm works well in the late stage.

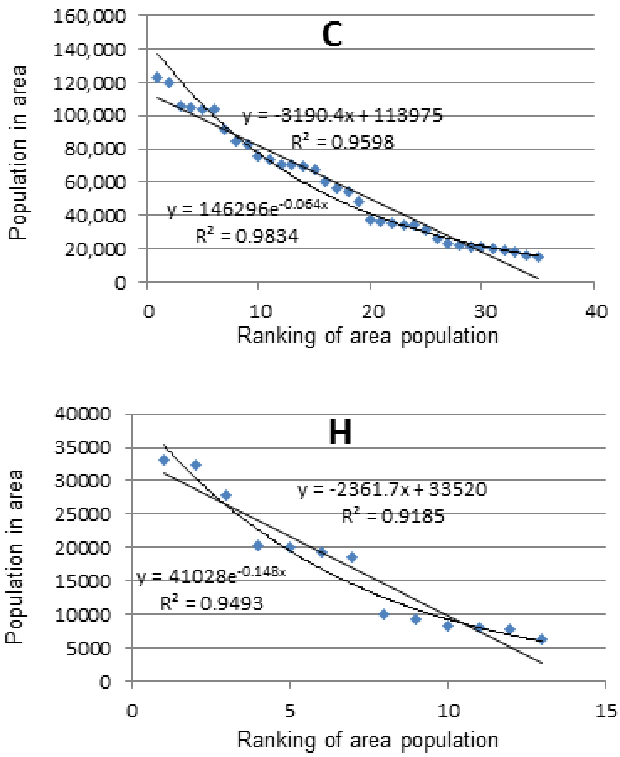

Figure 14. Relationship between population in the area and its ranking (Group 3).

The scatter diagrams in D, I, and J (Group 4) are shown in Fig. 15. The result showed that the approximation by exponential regression also fits into it at the highest R^2 = 0.9753 in D, while those by logarithmic regression fit into them at the highest R^2 = 0.9097, 0.9834 in I and J, respectively. Their population is considered to be linearly decreasing according to the regression line, since the R^2 difference of two regressions is small with the given microareas. Therefore, the effect by population difference is large because of linearity. Thus, the population-based algorithm works well in the late stage.

The scatter diagrams in B and F (Group 2) are shown in Fig. 16. The results showed that the approximation by logarithmic regression fits into them at the highest R^2 = 0.9424, 0.8435 in B and F, respectively. However, the difference of two regressions is large, linear regression does not fit in B or F. It is understandable that the population difference is small for low-ranking microareas in B and F according to logarithmic regression, so the effect of fluidity is greater than that of population.

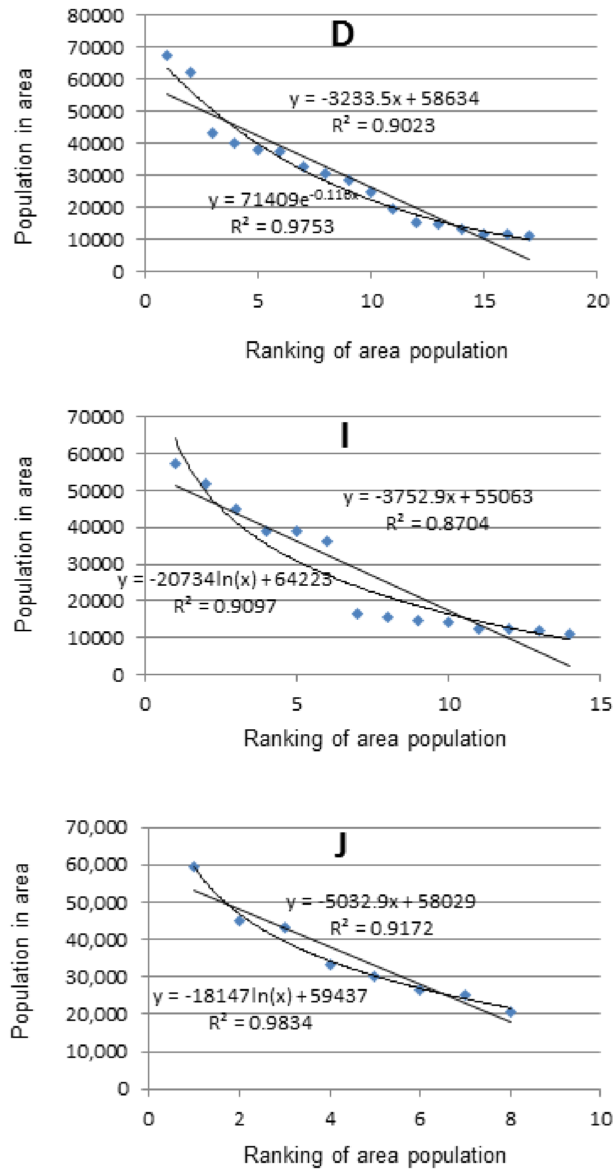

Figure 15. Relationship between population in the area and its ranking (Group 4).

The scatter diagrams in K, L, and M (Group 5) are shown in Fig. 17. The results showed that the approximation by logarithmic regression fit into them at the highest R^2 = 0.9645, 0.9513, 0.955, in K, L and M, respectively. However, the difference of two regressions is large, linear regression does not fit in K, L, or M. It is understandable that the population difference is small for low-ranking microareas in

K, L, and M according to logarithmic regression, so the effect of fluidity is greater than that of population.

The results with effect by fluidity are significant for Group 5, while the same results are obtained for Group 2. Note that the effect by fluidity and that by population are almost the same in Group 2 because of obtaining the same CR by the three algorithms.

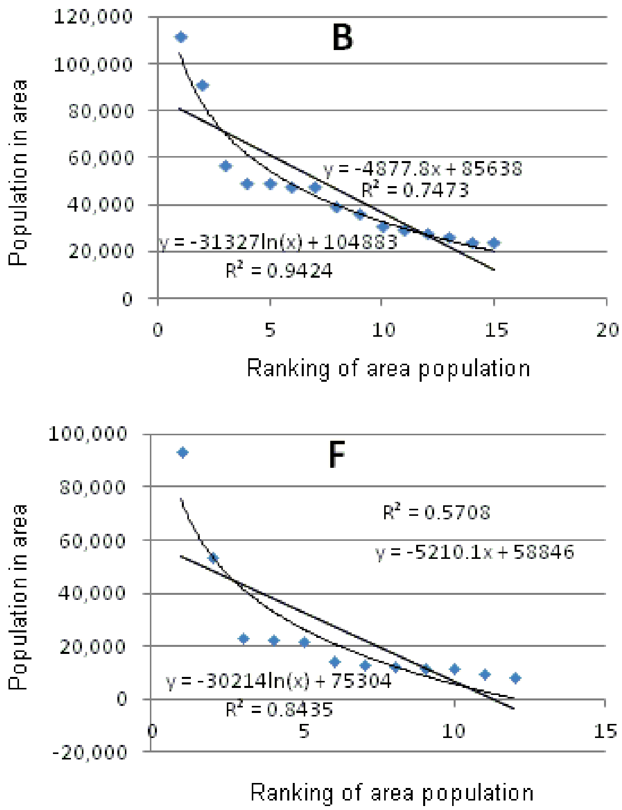

Figure 16. Relationship between population in the area and its ranking (Group 2).

The R^2 difference ratio (DR) is calculated using the following equation.

DR = 100 * {(R^2 of optimal regression) – (R^2 of linear regression)} /(R^2 of linear regression). (2)

The calculated results are shown in Table IV. DRs in Group 1, 3 and 4 are small, less than 10%. It means that the population with low-ranking microareas has linearity. Therefore, population difference has a large effect for microarea selection. Group 1 shows feature with large inflow. Group 3 shows features with large outflow. In addition, Group 4 has the feature of staying in its microarea as there is little movement. These groups are affected by population in the late stage.

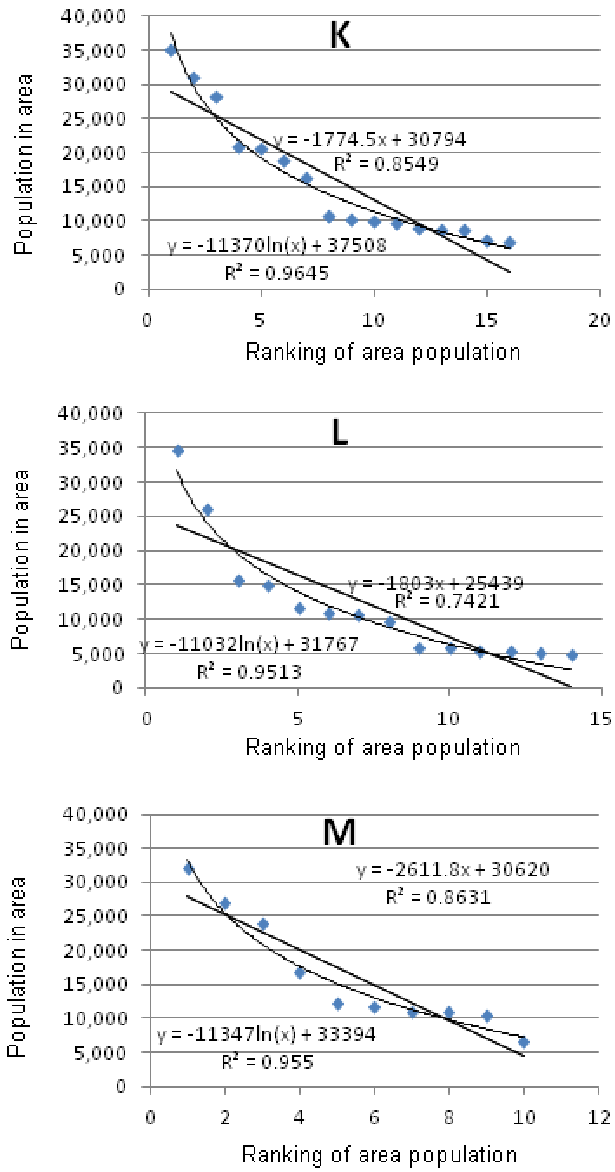

Figure 17. Relationship between population in the area and its ranking (Group 5).

DRs in Group 2 and 5 are large, greater than 10%. It means that the population with low-ranking microareas does not have linearity. Therefore, population difference has little effect for microarea selection. Group 2 has the feature with large in- and outflows as balanced movement, while Group 5 has the feature with small in- and outflows as balanced movement. These groups are affected by the fluidity even in the late stage.

VII. MICROAREA SELECTION METHOD BASED ON AREA CHARACTERISTICS

Summing up the results obtained in the previous section, the characteristics of each prefecture is an important factor.

These characteristics obtained by correspondence analysis are grouped into two types.

Char-1: Prefecture with biased in-/out-flow, or immobilization

Char-2: Prefecture with balanced in and outflows

The framework of the mixed algorithm considering area characteristics is constructed as follows.

- Step 1: Let areas be grouped by the correspondence analysis as ‘number of employees in own microarea’, ‘inflow of employees among microareas’, and ‘outflow of employees among microareas’ for inputs.
- Step 2: If the given area belongs to Char-1, then mixed algorithm is used at the given threshold value, α, i.e., flow-based algorithm is applied when the penetration rate is equal to or smaller than α, and population-based algorithm is applied when the penetration rate is greater than α.
- Step 3: If the given area belongs to Char-2, then the mixed algorithm is used under α = 100%, i.e., flow-based algorithm is applied during the life cycle.

TABLE IV. R^2-DIFFERENCE BETWEEN OPTIMAL AND LINEAR REGRESSIONS

Prefecture (Area no.)	Group	R^2 value of optimal regression	R^2 value of linear regression	Difference ratio: DR (%)
A (83)	1	0.9921	0.9364	6
B (54)	2	0.9424	0.7473	26
F (32)	2	0.8435	0.5708	48
C (87)	3	0.9834	0.9598	2
H (42)	3	0.9493	0.9185	3
D (43)	4	0.9753	0.9023	8
I (32)	4	0.9097	0.8704	5
J (28)	4	0.9834	0.9172	7
K (38)	5	0.9645	0.8549	13
L (35)	5	0.9513	0.7421	28
M (27)	5	0.955	0.8631	11

VIII. CONCLUSIONS

It is significantly important to select an area in which an ICT infrastructure should be introduced so as to ensure quick and economic development of an advanced information society. Area selection strongly depends on the potential demand, and one of the main features of the ICT infrastructure is network externality; therefore, a great amount of time and labour is required to select specified areas from among a large number of candidate areas.

In this paper, we proposed an efficient area selection method based on a service diffusion model. We evaluated the method using real field data from 13 prefectures, and we obtained the application of flow-based and population-based algorithms during the life cycle of the services. It is also advised that the areas are classified into two types in terms of employee fluidity and the application of mixed algorithm deeply depends on the microarea characteristics for population differences.

We intend to apply the method to other information network infrastructures, such as FTTH, LTE, and energy management services for future works.

REFERENCES

[1] M. Iwashita, A. Inoue, T. Kurosawa, and K. Nishimatsu, “Microarea selection method based on services diffusion process for broadband services,” Proc. of The Eleventh International Multi-Conference on Computing in the Global Information Technology (ICCGI 2016), pp. 30-35, 2016.

[2] The Ministry of Information and Communications, White Paper Information and Communications in Japan, 2009.

[3] The Ministry of International Affairs and Communications of Japan, “Broadband Service Coverage Rate With Respect to the Total Number of Households,” [Online]. Available from: http://www.soumu.go.jp/soutsu/tohoku/hodo/h2501-03/images/0110b1006.pdf (accessed 24 May 2017).

[4] The Ministry of International Affairs and Communications of Japan, “Information and Communications in Japan,” White Paper, p. 174, 2014.

[5] K. Yonetaka, Fact of area marketing, Tokyo: Nikkei Pub. Inc., 2008.

[6] Y. Murayama and R. Shibazaki, GIS theory, Tokyo: Asakura Pub. Co. Ltd., 2008.

[7] T. Sakai et al., “Use of Web-GIS area marketing and its application to the local region,” Bulletin of Global Environment Research, Rissho Univ., vol. 6, pp. 125-130, 2004.

[8] T. Ida, Broadband economics, Tokyo: Nikkei Pub. Inc., 2007.

[9] H. Seya and M. Tsutsumi, Spatial Statistics, Tokyo: Asakura Pub. Co. Ltd., 2014.

[10] S. Mase and M. Tsutsumi, Spatial Data Modeling, Tokyo: Kyoritsu Pub., 2001.

[11] R. L. Goodrich, Applied statistical forecasting, Belmont: Business Forecast Systems Inc., 1992.

[12] T. Abe and T. Ueda, “Telephone revenue forecasting using state space model,” Trans. on IEICE, vol. J68-A, no. 5, pp. 437-443, 1985.

[13] H. Kawano, T. Takanaka, Y. Hiruta, and S. Shinomiya, “Study on teletraffic method for macro analysis and its evaluation,” Trans. on IEICE, vol. J82-B, no. 6, pp. 1107-1114, 1999.

[14] M. Iwashita, K. Nishimatsu, T. Kurosawa, and S. Shimogawa, “Broadband analysis for network building,” Rev. Socionetwork Strat., vol. 4, pp. 17-28, 2010.

[15] M. J. Benner and M. L. Tushman, “Exploitation, Exploration, and Process Management: The Productivity Dilemma Revisited,” The Academy of Management Review, vol. 28, no. 2, pp. 238-256, 2003.

[16] R. W. White, “Motivation Reconsidered: The concept of competence,” Psychol. Rev., vol. 66, pp. 297-333, 1959.

[17] E. Rogers, “Diffusion of innovation,” 5th ed. New York: Free Press, 2003.

[18] S. Nakamichi, “NTT Docomo and The Univ. of Tokyo Have Just Started Analysis of Population Movement Using Location Information Obtained by Wireless Base Station,” [Online]. Available from: http://techon.nikkeibp.co.jp/article/NEWS/20100915/185680/ (accesses 24 May 2017).

[19] M. Iwashita, “A consideration of an area classification method for ICT service diffusion,” Knowl. Intell. Inf. Eng. Syst., LNAI, 6883, pp. 256-264, 2011.

[20] M. Iwashita, A. Inoue, T. Kurosawa, and K. Nishimatsu, "Micro area selection framework for ICT infrastructure diffusion based on commuting flow," Int. J. Comp. Inf. Sci., vol. 13, no. 2, pp. 10-19, 2012.

[21] M. Iwashita, A. Inoue, T. Kurosawa, and K. Nishimatsu, "Efficient microarea selection algorithm for infrastructure installation of broadband services," Int. J. Comput. Intell. Stud., Inderscience publishers, Vol. 5, Nos 3/4, pp. 29-236, 2017.

[22] K. Ikeda, "The Diffusion of Innovation Through Word-Of-Mouth and Social Networks: A social Psychological Study Combining Snowball Surveys and Multi-Agent Simulation," Tokyo: Univ. of Tokyo Press, 2010.

[23] S. P. A. T. Dentsu, "How to Create an Atmosphere to Make You Feel Buying," Tokyo: Diamond Inc., 2007.

[24] H. Katahira, "New theory AIDEES like big tidal stream," Diamond Visionary, vol. 42, no. 7, pp. 34-37, 2006.

[25] Macromill, "Field Survey of Using Internet Video Site". [Online]. Available from: http://monitor.macromill.com/researchdata/20070701movie/index.html (accessed 24 May 2017).

[26] "The Institute for Information and Communications Policy, Research Study of Determinative Factors for Internet Usage and Reality of Use" Tokyo: The Institute for Information and Communications Policy, 2009.

[27] S. Shimogawa and M. Iwashita, "Method for analyzing sociodynamics of telecommunication-based services," Proc. of 9th IEEE/ACIS International Conference on Computer and Information Science, pp. 99-104, 2010.

[28] G. A. Moor, "Crossing the chasm: Marketing and selling high-tech goods to mainstream customers," New York: Harper Business, 1991.

[29] Y. Ohsawa, M. Matsumura, and K. Takahashi, "Resonance without Response: The Way of Topic Growth in Communication," Chance Discov. Real World Decis. Making, Stud. Comput. Intell., Vol. 30, pp. 155-165, 2006.