THE OPTIMAL ARRANGEMENT OF INFRARED BEACONS ON A ROAD NETWORK TO COLLECT VEHICLE TRAJECTORIES - PATTERN ANALYSIS USING SCHEMA THEORY

Ryota HORIGUCHI
Doctoral Researcher
Institute of Industrial Science,
University of Tokyo
4-6-1-Cw504 Komaba, Meguro, Tokyo
153-8505 Japan
Fax: +81-3-54526420
E-mail: poepoe@nishi.iis.u-tokyo.ac.jp

Masao KUWAHARA
Professor
Institute of Industrial Science,
University of Tokyo
4-6-1-Cw504 Komaba, Meguro, Tokyo
153-8505 Japan
Fax: +81-3-54526420
E-mail: kuwahara@nishi.iis.u-tokyo.ac.jp

Hirokazu AKAHANE
Professor
Department of Civil Engineering,
Chiba Institute of Technology
2-17-1 Tsudanuma, Narashino, Chiba
275 Japan
Fax: +81-474-780474
E-mail: akahane@ce.it-chiba.ac.jp

Abstract: This research purposes to estimate the optimal arrangement of less infrared beacons to identify the trajectories of traveling vehicles on a road network. The paper, at first, describes the "schema", which represents the essential pattern to be extracted under certain beacon arrangement. Subsequently, the estimation of the optimal arrangement of beacons can be described as a combination optimization problem to maximize the entropy of schema. To demonstrate the estimation of the optimal arrangement, a case study in Hamamatsu area is attached at the end of this paper.

Key Words: infrared beacon, VICS, ITS, combination optimization

1. INTRODUCTION

An infrared beacon used in VICS (Vehicle Information Control System) in Japan has the role as a media both to provide the traffic information (Down-Link) and to collect the identical communication with equipped vehicles (Up-Link). It is now strongly expected to derive the valuable information like traveling vehicle trajectories from the collected Up-Links, and to utilize it for transportation planning and management.

This research purposes to estimate the optimal arrangement of less infrared beacons to collect the trajectories of all traveling vehicles on a road network. Here, a trajectory does not mean a precise vehicle tracking on a road, but means a path accompanied with timestamps on each link along the path. Although, the 'optimal' arrangement can be discussed from various points of view, e.g. optimal to provide most up-to-date traffic information for drivers, vehicle trajectories might be most useful information, which can be aggregated into any time-dependent OD matrix, or into travel time of any sections. Drivers' route choice behavior models, required for dynamic route guidance service, can be analyzed based on collected trajectories as well.

By simply chaining Up-Links of each identified vehicle, we may estimate the vehicle trajectories. If beacons are very sparsely installed, the estimated trajectory has some uncertainty because there might be alternative paths connecting two adjacent Up-Links' beacons. If beacons are fully installed onto each link, all of the trajectories of the equipped vehicle will completely determined, but this leads overinvestment off course. There must be an optimal arrangement of the beacons that ensures some certainty of estimated trajectories and satisfies the limit of the number of beacons. When the network becomes complicated and containing various directions of traveling vehicles, however, it is getting harder to estimate the optimal arrangement of beacons only with "chaining" Up-Links.

The paper, at first, introduces road-vehicle communication in VICS, then describes the methodology to evaluate the quality of the estimation using "schema", which is the feature pattern to be extracted from available paths under certain beacon arrangement. The "schema" can be applied to the future extended Up-Link, which contains the recent history of traveling links. Subsequently the estimation is described as a combination optimization problem using "schema". This optimization problem is NP-complete, but may be solved using some efficient searching methods. To demonstrate the estimation of the optimal arrangement, a case study in Tokyo area will be included in this paper.

2. UP-LINK COMMUNICATION IN VICS

There are three types of media used in VICS; i.e. i) infrared beacons, ii) FM short range wave beacons, and iii) FM broadcasting wave. Only infrared beacons support two-way communication, Down-Link and Up-Link, between road and equipped vehicles. Down-Link provides various traffic information, such as congestion, parking site availability, road construction, etc., which are to be indicated on the monitor of in-vehicle navigation system.

On the other hand, Up-Link collects vehicle identical information containing the items in Table 1. An equipped vehicle receives a random identification number (ID) given by the first beacon to pass through, and keep the ID during the trip. When an equipped vehicle passing through a beacon, the vehicle returns Up-Link to the beacon, packing with the ID number, present timestamp, previous beacon and previous timestamp, etc. The Up-Links are gathered and stored into the database in VICS center.

Table 1 : Items included in Up-Link

Data Item
Random vehicle ID
Current beacon ID
Current beacon passing timestamp
Previous beacon ID
Previous prefecture code
Previous beacon passing timestamp
Starting beacon ID
Starting prefecture code
Starting beacon passing timestamp
Vehicle type code

The Up-Links of the same ID are paired between two adjacent beacons, and normally used to estimate travel time of the section by comparing two timestamps of the pair (Oda *et al.*, 1996). Another usage to estimate origin-destination (OD) traffic volume has been proposed (Mashiyama *et al.*, 1998). The estimation method bases on the expansion of the sample OD matrix obtained by chaining Up-Links of the same ID. They also proposed the method to interpolate miscommunication of Up-Link, then attempted to estimate trajectory of each vehicle. However, they put the assumption that drivers may take the shortest path when there are alternative paths between the adjacent beacons. The estimated trajectories, therefore, are still not complete.

3. VEHICLE TRAJECTORY IDENTIFICATION USING SCHEMA

Although the possibility of the usage of Up-Links has been figured out, the problem how arrange the limited number of beacons still remains in order to obtain the valuable information with sufficient quality. The criteria for the optimal arrangement of beacon will be strongly required, because it is unrealistic that all links in a network implement beacons. When the network becomes complicated and contains various directions of traveling vehicles, however, it becomes difficult to estimate the optimal arrangement of beacons only by heuristic searching approach.

This chapter introduces the schema of path as an indicator of the degree of trajectory identification. Schema is the common patterns found in the set of pattern vectors. For instance, pattern vectors (1, 0, 0, 1) and (1, 1, 0, 0) are said to contain a schema (1, *, 0, *). The '*' in the schema denotes 'can be either 0 or 1'. Hereafter, a simple network shown in Figure 1 is used for the explanation of evaluation method in this chapter. The network consists of eight links and contains one OD pair.

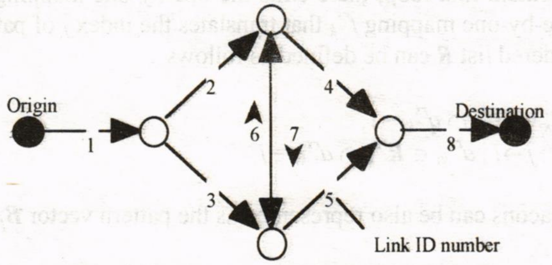

Figure 1 : Simple network used for the explanation

Before to start the evaluation, we need to enumerate all plausible paths found in the subject network. Some useful algorithms were proposed so far to enumerate plausible paths, e.g. the efficient path in Dial's assignment algorithm (Dial, 1970). In the sample network above, there are four paths between the OD pair.

The enumerated paths are represented in two ways. Firstly, the k-th path between an OD pair denoted as rs is expressed as an ordered list of link number $\boldsymbol{R}^{rs}_{k}$.

$$\boldsymbol{R}^{rs}_{k} = \{a^{rs}_{ki} \mid i = 1, 2, \ldots, m^{rs}_{k}\}. \tag{1}$$

where a^{rs}_{ki} : the i-th link number of the k-th path.

m^{rs}_k : the number of link along the k-th path.

In the sample network, the four paths are expressed as the following ordered lists, i.e.:

$$\begin{aligned} \boldsymbol{R}_1 &= \{1, 2, 4, 8\}, \\ \boldsymbol{R}_2 &= \{1, 3, 5, 8\}, \\ \boldsymbol{R}_3 &= \{1, 2, 7, 5, 8\}, \\ \boldsymbol{R}_4 &= \{1, 3, 6, 4, 8\}. \end{aligned} \tag{2}$$

The paths are also expressed as a pattern vector $\boldsymbol{Q}^{rs}_k$ of which elements take 0/1. The index of the vector means the link number. If the i-th link belongs to the k-th path, the i-th element of the vector takes 1.

$$\boldsymbol{Q}^{rs}_k = \{ q^{rs}_{kj} \,|\, q^{rs}_{kj} = \delta_j, j = 1, 2, \ldots, N\} \tag{3}$$

$$\begin{aligned} \text{where } \delta_j &= 1 \quad (a_j \in \boldsymbol{R}^{rs}_k), \text{ or} \\ &= 0 \quad (a_j \notin \boldsymbol{R}^{rs}_k), \end{aligned}$$

a_j : the j-th link,
N : the number of links on the k-th path.

Subsequently, the path $\boldsymbol{R}_1$~$\boldsymbol{R}_2$ are also expressed like the pattern vectors below.

$$\begin{aligned} \boldsymbol{Q}_1 &= (1, 1, 0, 1, 0, 0, 0, 1), \\ \boldsymbol{Q}_2 &= (1, 0, 1, 0, 1, 0, 0, 1), \\ \boldsymbol{Q}_3 &= (1, 1, 0, 0, 1, 0, 1, 0), \\ \boldsymbol{Q}_4 &= (1, 0, 1, 1, 0, 1, 0, 1). \end{aligned} \tag{4}$$

If any path does not contain link loop, there exits the one-by-one mapping between $\boldsymbol{R}^{rs}_k$ and $\boldsymbol{Q}^{rs}_k$. Therefore, the one-by-one mapping f^{rs}_k that translates the index j of path pattern vector $\boldsymbol{Q}$ to the index i of the ordered list $\boldsymbol{R}$ can be defined as follows.

$$\begin{aligned} &\forall\, q^{rs}_{kj} \in \boldsymbol{Q}^{rs}_k \cap q^{rs}_{kj} = 1, \\ &\exists\, f^{rs}_k : j \rightarrow i \,|\, a^{rs}_{ki} \in \boldsymbol{R}^{rs}_k \cap a^{rs}_{ki} = j. \end{aligned} \tag{5}$$

The arrangement of beacons can be also represented as the pattern vector $\boldsymbol{B}_p$, i.e.:

$$\boldsymbol{B}_p = \{b_{pl} \,|\, b_{pl} = \delta_l, l = 1, 2, \ldots, N\} \tag{6}$$

$$\begin{aligned} \text{where } \delta_l &= 1 \quad (\text{the } l\text{-th link has a beacon}), \text{ or} \\ &= 0 \quad (\text{the } l\text{-th link does not have a beacon}). \end{aligned}$$

Now, the operator '#' is introduced that extracts the schema $\boldsymbol{S}^{rs}_{kp}$ of the path $\boldsymbol{Q}^{rs}_k$ under the beacon arrangement $\boldsymbol{B}_p$.

$$\# : (\boldsymbol{Q}^{rs}_k, \boldsymbol{B}_p) \rightarrow \boldsymbol{S}^{rs}_{kp} \tag{7}$$

$$\begin{aligned} \text{where } \boldsymbol{S}^{rs}_{kp} &= \{s^{rs}_{kpl} \,|\, s^{rs}_{kpl} = \varphi_l, l = 1, 2, \ldots, N\} \\ \varphi_l &= 1 \quad (q^{rs}_{kl} = 1 \cap b_{pl} = 1), \\ &= 0 \quad (q^{rs}_{kl} = 0 \cap b_{pl} = 1), \text{ or} \end{aligned}$$

$$= \ * \qquad (b_{pl}=0).$$

Here '$s^{rs}_{kpl}=1$' implies that the path contains l-th link, and '$s^{rs}_{kpl}=0$' vice-versa. Otherwise, '$s^{rs}_{kpl}=*$' means unknown because there is no beacon installed at l-th link.

Assume that a vehicle traveling along one of the enumerated paths. The Up-Links concerning to the vehicle's trajectory can be expressed as an ordered list of link numbers, then also can be mapped into a link pattern vector, which contains a certain schema. This implies that the trajectory might be along one of the paths that embed the same schema within them.

Back to the sample network, assume the case that the two beacons are installed at Link-4 and Link-5. For this case, the pattern vector of beacon arrangement $\boldsymbol{B}_1$ is expressed as follows.

$$\boldsymbol{B}_1 = (0, 0, 0, 1, 1, 0, 0, 0) \tag{8}$$

Using the operator '#', the pattern vectors of the paths are transformed into the following schema vectors $\boldsymbol{S}_{11}\sim \boldsymbol{S}_{41}$.

$$\begin{aligned} &\boldsymbol{Q}_1\#\boldsymbol{B}_1 = \boldsymbol{S}_{11} = (*, *, *, 1, 0, *, *, *), \\ &\boldsymbol{Q}_2\#\boldsymbol{B}_1 = \boldsymbol{S}_{21} = (*, *, *, 0, 1, *, *, *), \\ &\boldsymbol{Q}_3\#\boldsymbol{B}_1 = \boldsymbol{S}_{31} = (*, *, *, 0, 1, *, *, *) = \boldsymbol{S}_{21}, \\ &\boldsymbol{Q}_4\#\boldsymbol{B}_1 = \boldsymbol{S}_{41} = (*, *, *, 1, 0, *, *, *) = \boldsymbol{S}_{11}. \end{aligned} \tag{9}$$

The schema vectors above are actually classified into two types $\boldsymbol{S}_{11}$ and $\boldsymbol{S}_{21}$, so that the Up-Links chain of all vehicle trajectories has one of $\boldsymbol{S}_{11}$ or $\boldsymbol{S}_{21}$. However, both $\boldsymbol{S}_{11}$ and $\boldsymbol{S}_{21}$ are associated with two paths. It is impossible to determine which path is the correct one in this case.

For the next case that the two beacons are installed at Link-2 and Link-4, the beacon pattern vector is described as follow.

$$\boldsymbol{B}_2 = (0, 1, 0, 1, 0, 0, 0, 0) \tag{10}$$

Similarly the following four schema vectors of the paths $\boldsymbol{R}_1\sim\boldsymbol{R}_2$ are obtained by applying the operator '#'.

$$\begin{aligned} &\boldsymbol{Q}_1\#\boldsymbol{B}_2 = \boldsymbol{S}_{12} = (*, 1, *, 1, *, *, *, *), \\ &\boldsymbol{Q}_2\#\boldsymbol{B}_2 = \boldsymbol{S}_{22} = (*, 0, *, 0, *, *, *, *), \\ &\boldsymbol{Q}_3\#\boldsymbol{B}_2 = \boldsymbol{S}_{32} = (*, 1, *, 0, *, *, *, *), \\ &\boldsymbol{Q}_4\#\boldsymbol{B}_2 = \boldsymbol{S}_{42} = (*, 0, *, 1, *, *, *, *). \end{aligned} \tag{11}$$

In this case, each of the four paths has its own schema, consequently any trajectory on the network can be identified to the path of which schema is the same as the Up-Links' schema.

4. OPTIMIZATION OF BEACON ARRANGEMENT

Getting the number of the paths that involve the same schema less, the certainty that the trajectory that gives the same schema would take one of the paths becomes higher. Therefore, by introducing the entropy of schemata, we may define the following optimization problem looking for the efficient arrangement of beacons to identify as much as vehicle trajectories with less beacons as possible.

$$\text{Find } \boldsymbol{B}_p \text{ such as } E(H_p, C_p) \rightarrow \text{maximum},$$

$$\text{where } H_p = -\Sigma_{rs}\Sigma_k x^{rs}_{kp}\log(x^{rs}_{kp}),$$
$$x^{rs}_{kp} = \|\boldsymbol{S}^{rs}_{kp}\| \,/\, K^{rs},$$
$$C_p = \|\boldsymbol{B}_p\| \,/\, N. \tag{12}$$

Here, H_p : the entropy of schemata,
C_p : the coverage ratio of the links with a beacon,
K^{rs} : the number of paths between rs,
$\|\boldsymbol{S}^{rs}_{kp}\|$: the number of paths belonging to $\boldsymbol{S}^{rs}_{kp}$,
$\|\boldsymbol{B}_p\|$: the number of links with a beacon,
N : the number of links.

The definition of entropy can be modified to incorporate realistic traffic conditions. For instance, if we prefer to identify the trajectories used by as much drivers as possible with less beacons, we may introduce the weighted entropy as follows.

$$H'_p = -\Sigma_{rs}Q^{rs}\Sigma_k x^{rs}_{kp}\log(x^{rs}_{kp}) \tag{13}$$

where Q^{rs} : traffic demand concerning with rs.

The entropy of schemata obviously becomes large when the number of the paths belonging to each schema is less. The arrangement giving the maximum entropy, therefore, implies that each schema has only one path, but such arrangement is not generally unique on the subject network. We have to decide the concrete definition of the evaluation function E both to maximize the entropy of schema and to minimize the number of beacons, depending on our purpose. For instance, if we want to completely identify the paths of all trajectories, E will be defined as follows; i.e.

$$E_1 = H_p + \delta(1 - C_p) \tag{14}$$

where $\delta = 1$ (when the all trajectories are identified), or
$= 0$ (otherwise)

It is clear that this optimization problem is NP-complete and has local minima, some effective searching algorithms to avoid the local minima need to be applied, such as genetic algorithms (Goldberg, 1989), TABU search method (Malek, 1990), etc.

5. SCHEMA FOR EXTENDED UP-LINK COMMUNICATION

It is now under consideration that the format of Up-Link is extended to include the nearest passing links and travel times. With this extended Up-Links, VICS may collect more information about the links having no beacon. We can expect to be able to identify vehicle trajectories with much less number of beacons than the present Up-Link format.

For the case of the extended Up-Link containing the history of n nearest links, the operator '#' in eq.(7) is modified to '$\#^n$' extracting schema $\boldsymbol{S}^{rs}_{kp}$ from link pattern vector $\boldsymbol{Q}^{rs}_k$ and beacon pattern vector $\boldsymbol{B}_p$; i.e.

$$\#^n : (\boldsymbol{Q}^{rs}_k, \boldsymbol{B}_p) \rightarrow \boldsymbol{S}^{rs}_{kp} \tag{15}$$

$$\begin{aligned} \text{where} \quad \boldsymbol{S}^{rs}_{kp} &= \{s^{rs}_{kpl} \mid s^{rs}_{kpl} = \varphi_l,\ l = 1, 2, \ldots, N\} \\ \varphi_l &= 1 \quad (\cup(q^{rs}_{ky}=1 \cap b_{py}=1 \mid (x=i\text{-}n, \ldots, i), f^{rs}_k(l) = i, f^{rs}_k(y) = x)), \\ &= 0 \quad (q^{rs}_{kl}=0 \cap b_{pl}=1), \text{ or} \\ &= * \quad (\text{otherwise}). \end{aligned}$$

Here, '$\varphi_l = 1$' means that the travel information of the l-th link would be included in the Up-Link uploaded within n further links.

Again back to the sample network and assume if Up-Link contains the history of 1 nearest link. When two beacons are installed at Link-4 and Link-5 (same as the pattern B_1 in eq.(8)), the schema of each path $\boldsymbol{R}_1 \sim \boldsymbol{R}_2$ is extracted as follows.

$$\begin{aligned} \boldsymbol{Q}_1\#^1\boldsymbol{B}_1 &= \boldsymbol{S}_{11} = (*, 1, *, 1, 0, *, *, *), \\ \boldsymbol{Q}_2\#^1\boldsymbol{B}_1 &= \boldsymbol{S}_{21} = (*, *, 1, 0, 1, *, *, *), \\ \boldsymbol{Q}_3\#^1\boldsymbol{B}_1 &= \boldsymbol{S}_{31} = (*, *, *, 0, 1, *, 1, *), \\ \boldsymbol{Q}_4\#^1\boldsymbol{B}_1 &= \boldsymbol{S}_{41} = (*, *, *, 1, 0, 1, *, *). \end{aligned} \tag{16}$$

In this case, four individual schemas can be extracted from four paths, and consequently all trajectories on the network can be identified.

6. COMPUTATIONAL EXPERIMENTS

In order to demonstrate the optimization of the beacon arrangement, the network at the center of Hamamatsu-city in Figure 2 is used for the computational experiments. The network spreads about 1.5 km from east to west, and contains 78 links. The 15 circles in the figure represent the centroids to be trip ends, so that there are 240 OD pairs in the network. Furthermore, 729 efficient paths can be enumerated.

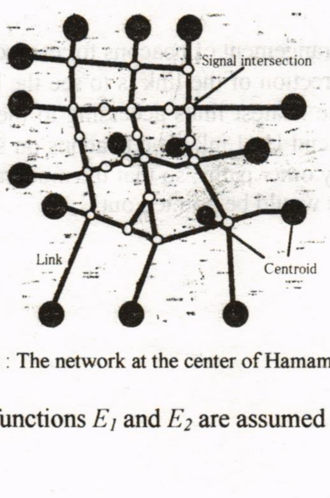

Figure 2 : The network at the center of Hamamatsu-city

Now, two types of evaluation functions E_1 and E_2 are assumed as follows; i.e.

$$E_1 = H_p + \delta(1 - C_p), \quad (17)$$
where δ = 1 (when the all trajectories are identified), or
= 0 (otherwise).

$$E_2 = H_p * (1 - C_p). \quad (18)$$

The eq.(17) implies to identify all trajectories with less beacons, and eq.(18) does compromise the identification and the number of beacons.

Then the next four case studies are examined. The result of the optimization for each case is indicated in Table 2. The optimization strategy here is somehow simple to repeat 1,000 times random Hill-Climbing then to employ the best for each case, as it is not the purpose of this research to determine the most useful strategy.

i) Optimize E_1 for the present Up-Link (n=0),
ii) Optimize E_1 for the extended Up-Link with nearest 3 links (n=3),
iii) Optimize E_1 for the extended Up-Link with nearest 5 links (n=5),
iv) Optimize E_2 for the extended Up-Link with nearest 5 links (n=5).

Table 2 : The result of optimization search

Case	# of paths identified	# of link installed beacon
i) E_1, n=0	729 (100%)	50 (64%)
ii) E_1, n=3	729 (100%)	37 (47%)
iii) E_1, n=5	729 (100%)	33 (42%)
iv) E_2, n=5	501 (69%)	18 (23%)

In the cases i) ~ iii), the optimal arrangement can identify 100% of trajectories. The number of beacons is getting smaller when the number of nearest links included extended Up-Link becomes large. However, the decrease in the number of beacon (50 -> 33) seems to be moderate comparing to the increase in the number of links in extended Up-Link (0 -> 5).

The compromising case of iv) identifies 69% of paths with only 23% coverage of beacon links. For this case, all paths are no weighted so that the path used by much trajectories and the path used by less trajectories are equally treated. If we consider the weight for each path depending on the OD traffic volume of the path, the percentage of identified trajectories will be get higher.

Figure 3 illustrates the optimal arrangement of beacons for each case. The arrows in the figure indicate beacons. The forward direction of the link is to see the beacon on the left-hand side. There may be a tendency that the nearest links accessing to the destination centroid always have beacons, no matter how extend Up-Link. The beacons on such accessing links may not be simultaneously used to identify other paths, so that the moderate decrease in the number of required beacons described above would be pointed out.

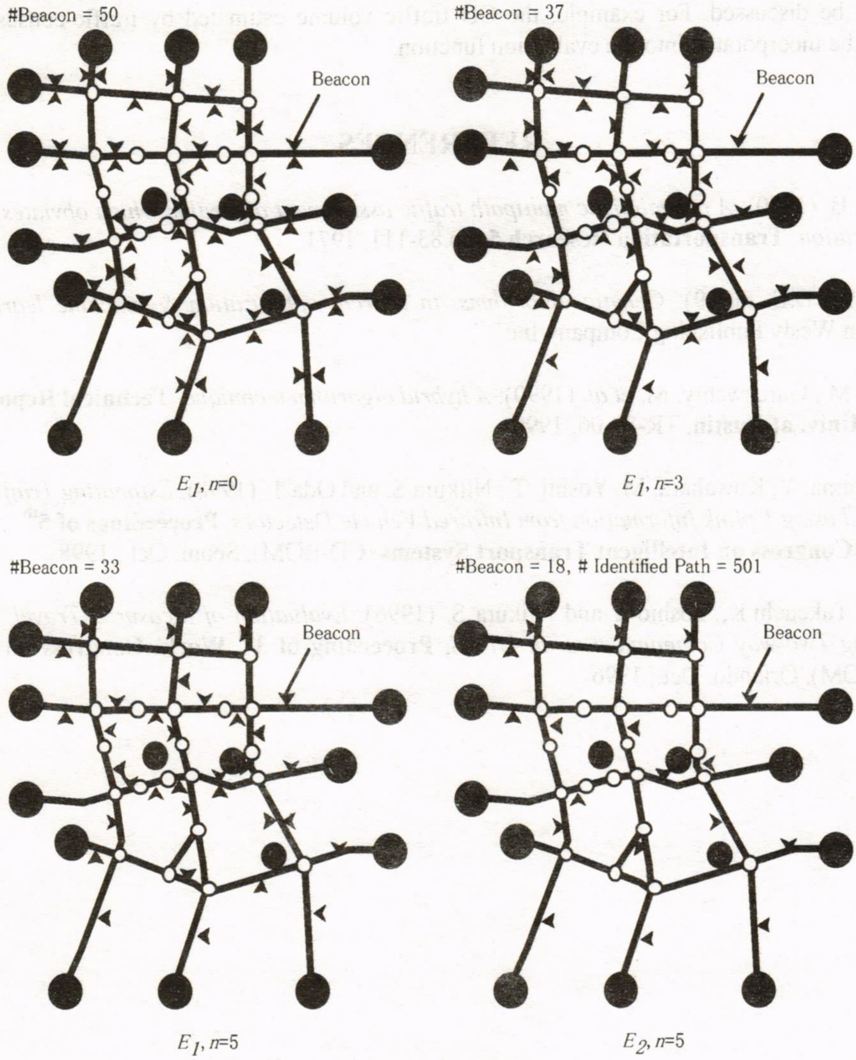

Figure 3 : The optimal arrangement of beacons

7. CONCLUSION

In this paper, the schema of the trajectory obtained under certain beacon arrangement on a road network was introduced as the index of path identification. Using the schema, it becomes possible to search the optimal arrangement of beacons to identify vehicle trajectories, such that it is difficult to obtain by heuristic method to chain Up-Link. The method also extended to future Up-Link format currently in consideration. Furthermore, the case studies using the real network at Hamamatsu-city were examined to demonstrate the applicability of the method, and the optimal arrangements of beacons were obtained.

For the practical use of the method presented here, there might be some discussion to be addressed. The first one coming is how to enumerate the plausible paths in advance. Dial's efficient paths are enumerated in this paper, but there is no guarantee that they are enough to cover the all paths drivers really use. What is the practical evaluation function for each case is also to be discussed. For example, the OD traffic volume estimated by traffic-census etc. should be incorporated into the evaluation function.

REFERENCES

Dial, R.B. (1970): *A probabilistic multipath traffic assignment algorithm which obviates path enumeration*, **Transportation Research 5**, pp.83-111, 1971

Goldberg, D.E. (1989): *Genetic algorithms, in search, optimization & machine learning*, Addison Wesly Publishing Company Inc.

Malek, M., Guruswamy, M. *et al.* (1990): *A hybrid algorithm technique*, **Technical Report of Texas Univ. at Austin**, TR-89-06, 1990

Mashiyama. Y, Kuwahara. M, Yoshii. T., Niikura S. and Oda T. (1998): *Estimating Traffic Demand using Uplink Information from Infrared Vehicle Detectors*, **Proceedings of 5th World Congress on Intelligent Transport Systems** (CD-ROM), Seoul, Oct., 1998

Oda.T, Takeuchi.K, Yoshio Y.and Niikura S. (1996): *Evaluation of Measured Travel Time Utilizing Two-way Communication in UTMS*, **Proceeding of 3rd World Congress on ITS** (CD-ROM), Orlando, Oct., 1996