The Theory of Traffic Assignment by Maximizing the Probability

Hiroshi MATSUI

Department of Civil Engineering
(Received September 13, 1971)

In this paper a new traffic assignment technique called *Probability Maximizing Method* is presented for the purpose of the theoretical descriptions of the realistic traffic assignment pattern on a real road network. To do this, the most probable assignment pattern under certain network and traffic conditions is considered.

This assignment model also makes it possible to use a travel time function (or travel time-volume relationship) which allows a continuous adjustment of link travel times as traffic volumes on the links increase.

Finally, this model is demonstrated by assigning hypothetical traffic volumes to a hypothetical network.

1. Introduction

It is well known that drivers' choices among various alternative routes through a road network as they travel from some origin to destination is very various and individual in general. On the other hand, it is also well known that traffic flow as a collective of these individual vehicles presents the assignment pattern wherein the shortest route is given the highest probability of use, on which we may recognize a certain statistical regularity. So it is reasonable to attempt to formulate a theory in which the traffic assignment pattern is derived from some equilibrium conditions.

A probabilistic model for estimating the trip distributions and assignments has been already developed in the previous works.[1),2)]

This paper extends such analyses to the application for the problem of assigning traffic to a road network using the capacity restraints procedure for a more realistic assignment.

2. Formulation of the Problem

To facilitate the formulation of the problem we consider a network of n nodes and l links, the former being partitioned in r nodes of trip origin, from which U_i $(i=1, 2, \cdots\cdots, r)$ trips begin, s nodes of trip destination, in which V_j $(j=1, 2, \cdots\cdots, s)$ trips end, and $(n\text{-}r\text{-}s)$ intermediate nodes. In the network we also consider q paths at most between any origin and destination node. A path is that series of links which constitutes the route through the network for an internodal trip. Finding these paths between each OD pair is usually very cumbersome and time consuming for a large network. To date, however, some efficient algorithms for determination of the paths through the network have been already suggested.[3)]

Consider X trips distributed on the network on which the total travel time is E. Seeing these trips by distinguishing them individually, there are many micro-states, or combinations of trips. By formula of combinations the total number of such micro-states $W_X(E)$ can be given by

$$W_X(E)=\frac{(E+X-1)!}{(X-1)!E!} \qquad (1)$$

where E is taken to be an integer value.[4]

Next let $W_X(E+\varDelta E)$ be the total number of micro-states when E changes infinitestimally $\varDelta E$, then it can be written as

$$W_X(E+\varDelta E)=\frac{(E+\varDelta E+X-1)!}{(X-1)!(E+\varDelta E)!} \qquad (2)$$

Taking the ratio of Eq. (2) to Eq. (1), we have

$$\begin{aligned} W_X(E+\varDelta E)/W_X(E) &= (E+\varDelta E+X-1)!(X-1)!E!/(E+X-1)!(X-1)!(E+\varDelta E)! \\ &= (E+X)(1+X/(E+1))(1+X/(E+2))\cdots\cdots \\ &\quad \cdots\cdots(1+X/(E+\varDelta E-1))/(E+\varDelta E) \end{aligned} \qquad (3)$$

Assuming $X\gg 1$ and $E\gg\varDelta E$ we omit $\varDelta E$ in Eq, (3), we have

$$\frac{W_X(E+\varDelta E)}{W_X(E)}=(1+1/\bar{t})^{\varDelta E} \qquad (4)$$

where we put $E/X=\bar{t}$. $\bar{t}$ means the average travel time per trip.

Taking the logarithm of both sides of the above equation, we have

$$\log W_X(E+\varDelta E)-\log W_X(E)=\varDelta E \log (1+1/\bar{t}) \qquad (5)$$

Then

$$\varDelta \log W_X(E)=\gamma\varDelta E \qquad (6)$$

where we put

$$\gamma=\log (1+1/\bar{t}) \qquad (7)$$

Next consider the number of micro-states z_E corresponding to a certain distribution $\{X_{ij}^k\}$ $(i=1, 2, \cdots\cdots, r,\ j=1, 2, \cdots\cdots, s,\ k=1, 2, \cdots\cdots, q)$, then it can be expressed by

$$z_E=\left(\frac{X!}{\prod\limits_{ijk} X_{ij}^k!}\right)_E \qquad (8)$$

from formula of combinations. Where X_{ij}^k is the number of trips from origin node i to destination node j via the k^{th} available path and a set $\{X_{ij}^k\}$ should be taken within the limits of the total travel time being E.

Consequently the previous function (1) can be expressed by the sum of these z_E in all combinations of various X_{ij}^k, that is

$$W_X(E)=\sum_{X_{ij}^k}\left(\frac{X!}{\prod\limits_{ijk} X_{ij}^k!}\right)_E \qquad (9)$$

It is proved mathematically that assuming that all micro-states are equally probable and X is large enough $W_X(E)$ can be replaced by the maximum of z_E for almost certain. Hence the trip distribution $\{X_{ij}^k\}$ for which z_E is a maximum can be shown to be overwhelmingly the most probable. So we designate the maximum of z_E by Z_E, the following relaion holds

$$\frac{W_X(E+\varDelta E)}{W_X(E)}=\frac{Z_{E+\varDelta E}}{Z_E} \qquad (10)$$

From Eq. (6) and (10)

$$\varDelta \log W_X(E)=\varDelta \log Z_E=\gamma\varDelta E \qquad (11)$$

Subtracting $\varDelta \log z_E$ from both sides of the latter half relation of Eq. (11) we have

$$\varDelta (\log Z_E-\log z_E)=\varDelta (\gamma E-\log z_E) \qquad (12)$$

Hence the most probable trip distribution and assignment pattern is given by minimizing

$$\gamma E - \log\ z_E \qquad (13)$$

It is noted that Eq. (13) is homologous to the Helmholtz free-energy function which has a minimum for a thermodynamic system in equilibrium under the conditions of constant volume and temperature and log z_E of Eq. (13) is analogous to a quantity defined to be entropy in statistical mechanics.

Next we denote the notation as follows:

u_i=normalized supply of trips in origin node i (i=1, 2, ……, r),

v_i=normalized demand of trips in destination node j (j=1, 2, ……, s),

P_{ij}=transition probability of a trip from origin node i to destination node j,

p_{ij}^k=probability of a trip from origin node i to destination node j via the k^{th} available path (k=1, 2, ……, q).

Using the above notation X_{ij}^k can be written by

$$X_{ij}^k = X u_i P_{ij} p_{ij}^k \qquad (14)$$

Next, as we have the travel time t_{ij}^k required from origin node i to destination node j via the k^{th} path, the total travel time E on the whole trips X is written by

$$E = \sum_i\sum_j\sum_k X_{ij}^k t_{ij}^k = X\sum_i\sum_j\sum_k u_i P_{ij} p_{ij}^k t_{ij}^k \qquad (15)$$

Similarly the second term of Eq. (13), which we may call entropy function owing to the form of the Helmholtz free-energy function, can be written by

$$\log\ z_E = -X\sum_i\sum_j\sum_k u_i P_{ij} p_{ij}^k \log\ p_{ij}^k - X\sum_i\sum_j u_i P_{ij} \log\ P_{ij} - X\sum_i u_i \log\ u_i \qquad (16)$$

herein Stirling's formula ($\log\ x! \doteqdot x\ \log\ x - x$) in used.

If OD table is given, that is, if u_i, v_j and P_{ij} are constants, the traffic assignment problem is to choose p_{ij}^k to

$$\text{minimize}\ (\gamma\sum_i\sum_j\sum_k u_i P_{ij} p_{ij}^k t_{ij}^k + \sum_i\sum_j\sum_k u_i P_{ij} p_{ij}^k \log\ p_{ij}^k) \qquad (17)$$

subject to

$$\sum_k p_{ij}^k = 1 \qquad (18)$$

Solutions can be obtained by Lagrangian method in general.

Assuming that $t_{ij}^k s$ are constant without regard to traffic volumes, we have

$$p_{ij}^k = \frac{\exp(-\gamma t_{ij}^k)}{\sum_k \exp(-\gamma t_{ij}^k)} \qquad (19)$$

Hence, in this case OD traffic volumes are assigned to the respective paths by Eq. (19) and link volumes are simply accumulated without regard to link capacities.

It is well known that delay caused by congestion will persuade some travelers to change their destination entirely. Such a formulation that both trip distributions and assignments are solved simultaneously by giving trip generation and attraction in each node is more suitable for this problem.

The problem is to choose P_{ij} and p_{ij}^k to

$$\text{minimize}\ (\gamma\sum_i\sum_j\sum_k u_i P_{ij} p_{ij}^k t_{ij}^k + \sum_i\sum_j\sum_k u_i P_{ij} p_{ij}^k \log\ p_{ij}^k + \sum_i\sum_j u_i P_{ij} \log\ P_{ij}) \qquad (20)$$

subject to

$$\sum_j P_{ij} = 1 \qquad (21)$$

and

$$\sum_i u_i P_{ij} = v_j \qquad (22)$$

For p_{ij}^k we have the same form of solution as in Eq. (19) and for P_{ij} we have

$$P_{ij} = \alpha_i \beta_j \exp(-1 - \sum_k p_{ij}^k \log\ p_{ij}^k - \gamma\sum_k p_{ij}^k t_{ij}^k) \qquad (23)$$

where,

$$\alpha_i = e/\sum_j \beta_j exp\,(-\sum_k p_{ij}^k\ log\ p_{ij}^k - \gamma\sum_k p_{ij}^k t_{ij}^k) \cdots\cdots (24)$$

and

$$\beta_j = ev_j/\sum_i \alpha_i u_i exp\,(-\sum_k p_{ij}^k\ log\ p_{ij}^k - \gamma\sum_k p_{ij}^k t_{ij}^k) \cdots\cdots (25)$$

herein we put $\alpha_i = exp\,(-\mu_i/u_i)$ and $\beta_j = exp\,(-\varphi_j)$. μ_i and φ_j are Lagrange multipliers associated with Eq. (21) and (22).

P_{ij} can be obtained by the following iterative procedure.

1) Assume γ and calculate a set $\{p_{ij}^k\}$ from Eq. (13).
2) Assuming a set $\{\beta_j\}$ and using the above p_{ij}^k values calculate a set $\{\alpha_i\}$ from Eq. (24).
3) Calculate a set $\{\beta_j\}$ by substituting the above α_i values into Eq. (25).
4) Return to step 2) and repeat the procedure until the new $\{\alpha_i\}$ and $\{\beta_j\}$ are sufficiently close to the previous sets to indicate adequate convergence respectively.
5) Calculate a set $\{P_{ij}\}$ using the converged $\{\alpha_i\}$ and $\{\beta_j\}$ from Eq. (23).

In the above procedure we may determine the value of γ in order that the calculated assignment pattern may be suited to the actual one. Particularly, when the total travel time (or the average travel time per trip) is given, we may determine the value of γ by the iterations in order that the calculated total travel time may be sufficiently close to the actual one.

3. Traffic Assignment with a Travel Time Function

The traffic assignment process described above does not take the effect of link capacity on traffic flow into account, and consequently many links in the network become unrealistically oveloaded. Therefore the necessity of introducing link capacity restraints - the cause of congestion is realized.

The primary feature of the capacity-restraint algorithm is its provision for arbitary link travel time functions. These functions may be any linear or nonlinear relation of link flow to travel time for that flows.

Let us redefine the total travel time E using link travel time function, it is given by

$$E = X\sum_i\sum_j\sum_k\sum_h {}_{ijk}\delta_h u_i P_{ij} p_{ij}^k f_h\ (\sum_i\sum_j\sum_k {}_{ijk}\delta_h u_i P_{ij} p_{ij}^k) \cdots\cdots (26)$$

where

$${}_{ijk}\delta_h = \begin{cases} 1 & \text{if link } h \in \text{path } k \text{ between node } i \text{ to } j. \\ 0 & \text{if link } h \notin \text{path } k \text{ between node } i \text{ to } j. \end{cases}$$

and $f_h\ (\sum_i\sum_j\sum_k {}_{ijk}\delta_h u_i P_{ij} p_{ij}^k)$ is the travel time function on link h related to the link volume.

In this case the assignment problem is to find a set of p_{ij}^k to

$$\text{minimize}\ [\gamma\sum_i\sum_j\sum_k\sum_h {}_{ijk}\delta_h u_i P_{ij} p_{ij}^k f_h\ (\sum_i\sum_j\sum_k {}_{ijk}\delta_h u_i P_{ij} p_{ij}^k) + \sum_i\sum_j\sum_k u_i P_{ij} p_{ij}^k log p_{ij}^k] \cdots (27)$$

subject to

$$\sum_k p_{ij}^k = 1 \cdots\cdots (28)$$

It is proved that if f_h is convex and increases monotonously the function (27) is a convex function and has only one minimum value. This proves the uniqueness of the solution for this assignment problem.

The following iterative procedure is used for obtaining a set of p_{ij}^k.

1) Assume γ and calculate a set $\{p_{ij}^k\}$ based on the path travel time $t_{ij}^k(1)$ between

each OD pair at free speed.

2) Using $\{p_{ij}^k\}$ calculated above OD volumes are assigned to the respective paths and link volumes are accumulated.

3) The link travel time function are utilized to revise the link travel times.

4) New path travel times $t_{ij}^k(1)'$ between each OD pair is calculated based on the revised link travel times. The revised path travel times for the second and succeeding iterations are updated according to

$$t_{ij}^k(n)=\frac{mt_{ij}^k(n-1)+t_{ij}^k(n-1)'}{m+1} \quad n=2,\ 3,\ \cdots\cdots \tag{29}$$

where m is a constant introduced in order to stabilize these iterations.

5) Return to step 1) these iterations are repeated using $t_{ij}^k(n)$ in place of $t_{ij}^k(n-1)$ until $t_{ij}^k(n)$ is sufficiently close to $t_{ij}^k(n-1)$ to indicate adequate convergence.

6) Based on the converged path travel times OD volumes are assigned and link volumes are accumulated.

In the above iterations the value of m may be usually 1. However, when we use the link travel time function such that the change in travel time as flow increases is small for low flow values, but very large as saturation flow is approached (e. g. the logarithmic travel time function used in the next example problem), these iterations may tend to oscillate without convergence. In such a case we had better enlarge the value of m properly in order to stabilize these iterations.

4. An Example Problem

To demonstrate the use of this traffic assignment method an example problem was solved. In **Fig. 1** the problem network is shown. The trip distribution pattern composed of six OD pairs is also given in **Table I**. Three paths are considered between each OD pair as in **Fig. 2, 3** and **4**. These paths are fixed during the assignment process.

i) Traffic Assignment with a Constant Travel Time

In this case OD volumes are assigned to the respective paths by Eq. (19) and link volumes are simply accumulated. The value of γ was determined by an iterative procedure so that the average travel time per trip might become 8.5 minutes.

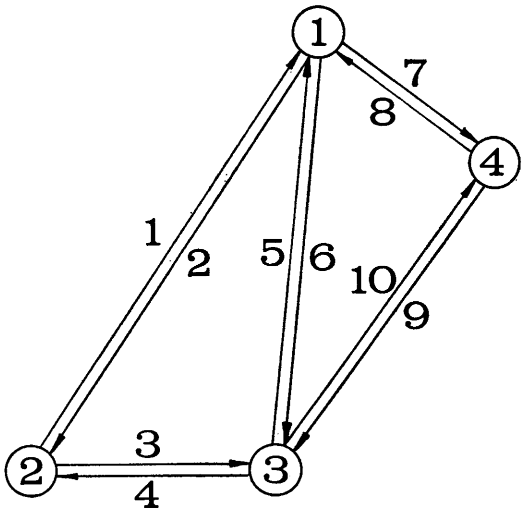

Fig. 1 Road network for the example problem

Table I OD table (vehicles per day)

O \ D	1	2	3	4
1	*	2,600	1,700	*
2	2,300	*	*	700
3	1,500	*	*	*
4	*	1,200	*	*

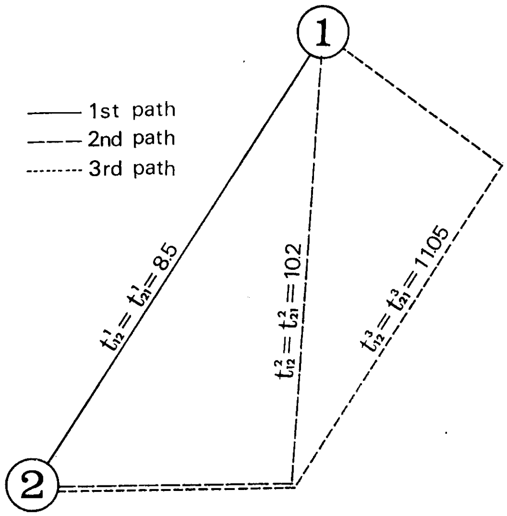

Fig. 2 Set of paths between Node 1 and 2 and path travel times (minutes)

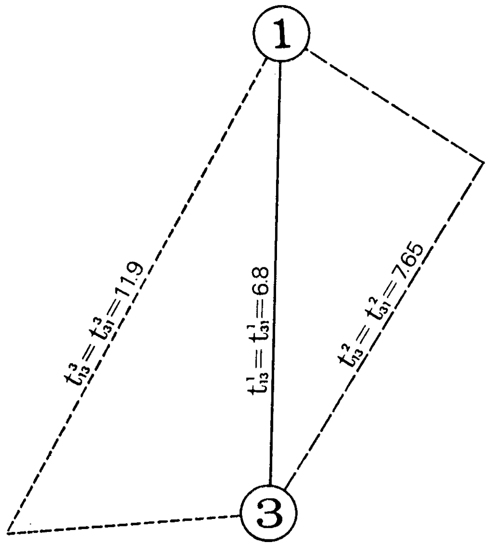

Fig. 3 Set of paths between Node 1 and 3 and path travel times (minutes)

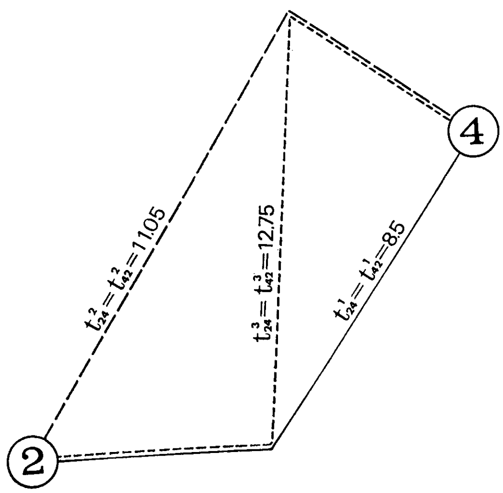

Fig. 4 Set of paths between Node 2 and 4 and path travel times (minutes)

The result of five iterations was that γ was 0.6949.

The assigned and link volumes are shown respectively in **Table Ⅱ** and **Fig. 5.** To make a comparison with the results of this traffic assignment pattern two extreme assignment techniques were used to assign on a common network and trip table. One is an all-or-nothing assignment which is brought when we put $\gamma=\infty$ in Eq. (17). The other is an uniform path assignment, which is brought when we put $\gamma=0$ in Eq. (17). **Fig. 6** indicates the result of the all-or-nothing assignment pattern and **Fig. 7** the uniform path assignment pattern.

Table Ⅱ Assignment with a constant travel time (vehicles par day)

OD	1st path	2nd path	3rd path
X_{12}	1,761	540	299
X_{13}	1,074	595	31
X_{21}	1,557	478	265
X_{24}	573	97	30
X_{31}	948	525	27
X_{42}	982	167	51

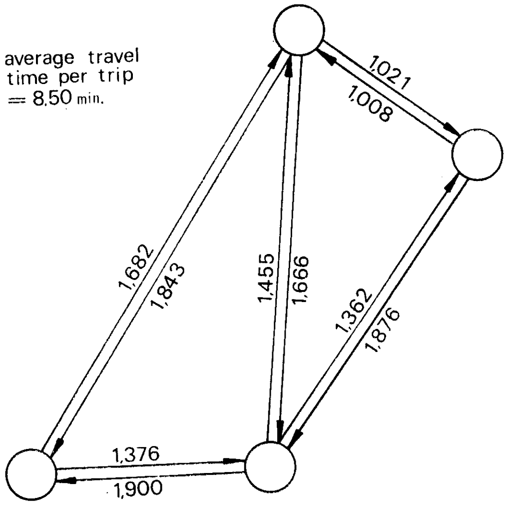

Fig. 5 Link volumes resulting from assignment with a constant travel time (vehicles per day) $\gamma=0.6949$

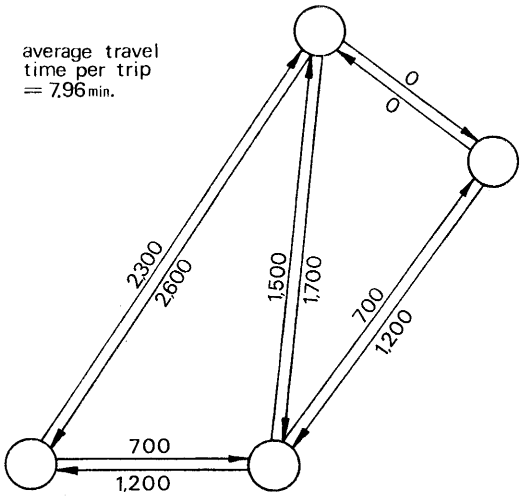

Fig. 6 Link volumes resulting from assignment on an all-or-nothing basis (vehicles per day) $\gamma=\infty$

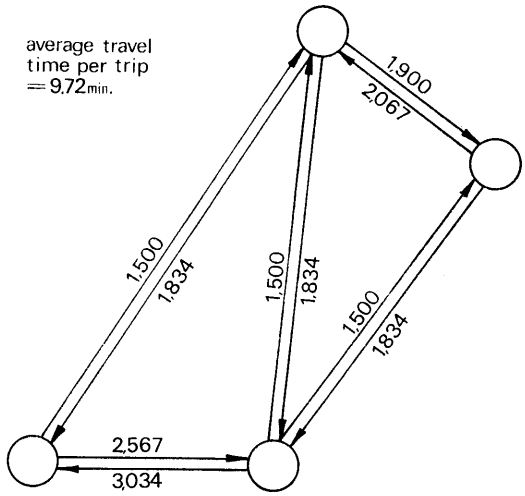

Fig. 7 Link volumes resulting from assigning traffic uniformly to each path (vehicles per day) $\gamma=0$

ii) Traffic Assignment with a Travel Time Function

The following two travel time functions were employed for this example.

$$\text{Type 1} \quad T_h = A_h Q_h + B_h \quad \cdots\cdots (30)$$

$$\text{Type 2} \quad T_h = \begin{cases} A_h \log (2,000/2,000 - Q_h) + B_h & 0 \leq Q_h < 2,000 \\ \infty & 2,000 \leq Q_h \end{cases} \quad \cdots\cdots (31)$$

where
T_h=travel time on link h in minutes
Q_h=link volume on link h in vehicles per day.
A_h=empirically derived constant
B_h=constant representing link travel time at free flow conditions.

Type 1 is a linear travel time function. This function allows travel time to increase linearly with flow. While *Type 2* is a nonlinear, logarithmic travel time function, the characteristics of this function are such that the change in travel time as flow increases is small for low flow values, but large as saturation flow is approached. This function can prevent flows in excess of link capacity from being assigned to the links. **Fig.** 8 shows these travel time functions. **Table** Ⅲ and Ⅳ give A_h and B_h for each of the links.

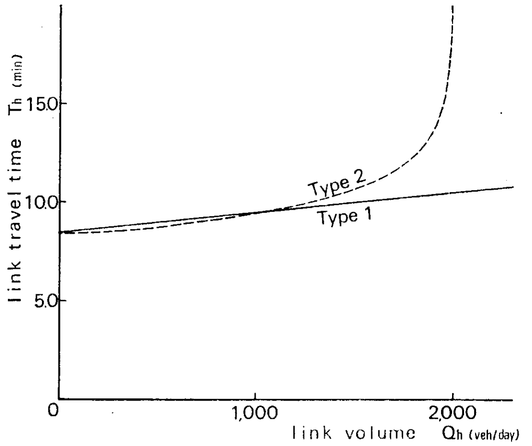

Fig. 8 Link travel time function (for Link 1)

Table Ⅲ Link travel time functions of *Type 1*

Link	A_h	B_h
1 and 2	0.0010	8.5
3 and 4	0.0004	3.4
5 and 6	0.0008	6.8
7 and 8	0.0003	2.55
9 and 10	0.0006	5.1

Table Ⅳ Link travel time functions of *Type 2*

Link	A_h	B_h
1 and 2	1.4427	8.5
3 and 4	0.5771	3.4
5 and 6	1.1542	6.8
7 and 8	0.4328	2.55
9 and 10	0.8656	5.1

The solutions to the problem with a linear travel time function are shown in **Fig. 9** and **Table** Ⅴ. To this problem the value of γ was determined so that the average travel time per trip might become 10.0 minutes and consequently γ was 0.6399.

Similarly the solutions to the problem with a logarithmic travel time function are shown in **Fig. 10** and **Table** Ⅵ. To this problem γ was determined so that the average travel time per trip might become 11.04 minutes and γ was 0.7291.

The value of m used in the iterations to obtain the solutions might be 1 for *Type 1*, but for *Type 2* it was changed in steps, that was, m=39 for iteration steps between first and 16 th, m=29 for steps between 17 th and 29 th and m=9 for 30 th and over. A more rational procedure for determining the value of m will reguire

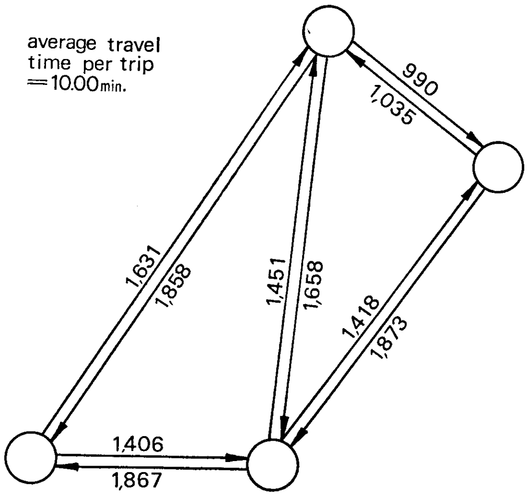

Fig. 9 Link volumes resulting from assignment with a travel time function of *Type 1* (vehicles per day) $\gamma=0.6399$

Table V Assignment with a travel time function of *Type 1* (vehicles per day)

OD	1st path	2nd path	3rd path
X_{12}	1,788	525	287
X_{13}	1,086	593	21
X_{21}	1,531	486	282
X_{24}	591	83	26
X_{31}	939	545	16
X_{42}	992	161	47

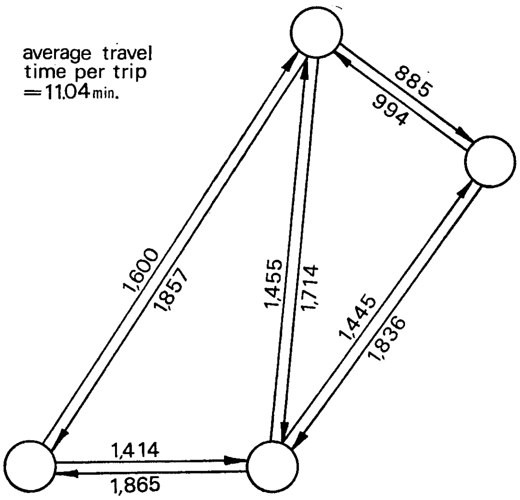

Fig. 10 Link volumes resulting from assignment with a travel time function of *Type 2* (vehicles per day) $\gamma=0.7291$

Table VI Assignment with a travel time function of *Type 2* (vehicles per day)

OD	1st path	2nd path	3rd path
X_{12}	1,811	533	256
X_{13}	1,143	549	8
X_{21}	1,533	487	280
X_{24}	620	61	19
X_{31}	949	545	6
X_{42}	1,031	131	38

further research.

Next, in order to know the relationship between γ and $\bar{t}$ many traffic assignments were calculated under the assumption of the various values of γ. Then we had the results as shown in **Fig. 11.**

It is noted that for the assignment with a constant travel time γ-$\bar{t}$ curve becomes a monotonous decreasing curve, while for the assignment with a travel time function it presents a curve that has a minimum value. Therefore it is apparent that two values of γ correspond to a single $\bar{t}$ when we use a travel time function. This will show that we must determine which we should choose of the two by observing the actual assignment pattern.

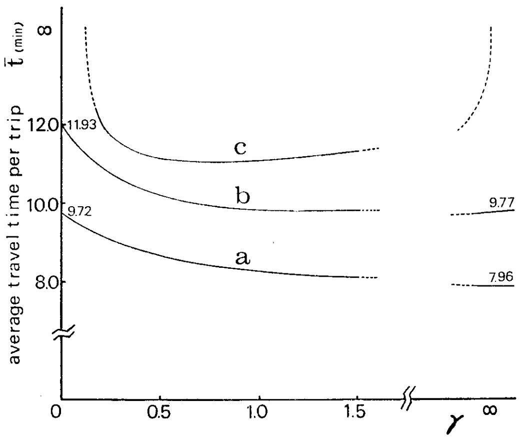

Fig. 11 Relationship between γ and $\bar{t}$
a: assignment with a constant travel time
b: assignment with a travel time function of *Type 1*
c: assignment with a travel time function of *Type 2*

5. Conclusion

Some traffic assignment technigues with a travel time function whereby trips are assigned to the paths specified beforehand have been suggested in the previous works.[6),7)] The assignment method proposed in this paper may be structually common with such techniques but different from the following characteristics.

i) It is necessary to specify some available paths beforehand between the points of origin and destination, but the assignment factor to each of these paths can be determined endogenously by this assignment method.

ii) Any linear or nonlinear travel time function may be applied to this method. Especially if a travel time function is convex and increases monotonously, the uniqueness of the solution can be proved mathematically.

iii) Not only trip assignments but also trip distributions can be estimated simultaneously by this method. On the trip distribution phase, the reader is referred to the previous paper[8)] for a detailed account.

iv) In the objective function (13), if γ is so large that the second term is negligible we have an optimal traffic assignment pattern such that the total travel time is minimized. Moreover, if link travel times are not related to link flows, it reduces to an all-or-nothing assignment pattern. While if γ is 0, the problem is to maximize entropy function, which means that OD volumes are assigned to the respective paths uniformly.

In this paper the most probable assignment pattern on the network was considered from a stochastic points of view. However whether the traffic assignment pattern by this method can be describe or not the actual one on the real network is another question. Accordingly for further development it is necessary to investigate the applicability of this technique to practical traffic assignment problems.

References

1). Matsui H.: "Trip Distribution Analysis Based on an Analogue of Statistical Mechanics", Preprint of the 24th Annual Conference of the Japan Soc. Civil Engrs. Part 4, pp. 207~208, September 1969.

2). Matsui H.: "Some Stochastic Aspects of Traffic Distribution Patterns", Proc. of the Japan Soc. Civil Engrs. No. 190, pp. 99~109, June 1971.

3). e. g. Moore E. F.: "The Shortest Path Through a Maze", Proc. Internat. Symp. on the Theory of Switching, April 1957.

4). Kubo R.: "Statistical Mechanics", Kyoritsu Publishing Co. October 1964.
5). e. g. Fast J. D.: "Entropy-the Significance of the Concept of Entropy and Its Applications in Science and Technology", Philips Technical Library, 1968.
6). Kato A.: "The Application of Traffic Flow Analysis for the Road Network Planning", Roads, September and October 1964.
7). Hoshino T.: "The Theory of Traffic Assignment on the Road Network", Annual Report of the Fifth Japan Roads Conference, pp. 68～70, October 1959.
8). Reference 2).