A FUZZY INFERENCE BASED ASSIGNMENT ALGORITHM TO ESTIMATE O-D MATRIX FROM LINK VOLUME COUNTS

K. Harikishan Reddy and Partha Chakroborty
Indian Institute of Technology, Kanpur 208 016, India

ABSTRACT. Origin destination trip demand matrix (O-D) is an essential ingredient in a wide variety of travel analysis and planning studies. The traditional way of obtaining O-D from home-interview survey data is expensive in terms of time and money. O-D estimation from link counts on the other hand is more appealing as the required data collection is simple and routine. Some of the O-D estimation methods which use link counts require the relative contributions of the O-D elements toward the link volumes. In the literature these proportions are assumed exogenous and are usually obtained from proportional traffic assignment models. In this study a flow dependent fuzzy inference based assignment algorithm is proposed for generating the required proportions and a recursive methodology is adopted for the O-D estimation. The methodology is applied to three test networks and the results are found to be satisfactory.

INTRODUCTION

The O-D is an essential ingredient in a wide variety of travel analysis and planning studies (Nguyen, 1984). The traditional way of estimating O-D from home-interview survey data is expensive (Bierlair & Toint, 1995). Hence, generally, the estimates are based on small sample of home-interview data and thus the accuracy of the estimates suffer. This led the researchers to estimate the O-D from a variety of other data sources among which O-D estimation from link traffic counts has attracted lot of interest as the required data collection is simple and routine. The currently available methods for O-D estimation and their merits and demerits are given in the next section. The problems with current methods prompted the development of the proposed approach which relies on a fuzzy inference based flow dependent assignment algorithm for generating the route choice proportions for the estimation of the O-D. The O-D is estimated using bilevel optimization.
LITERATURE REVIEW

O-D estimation from equilibrium assignment: The models in this class are based on the assumption that the O-D when assigned according to Wardrop’s first principle will reproduce the observed arc flows. Note that, this assumption necessitates that the observed flows be in equilibrium. Nguyen (1984) was the first to attempt the equilibrium assignment based O-D estimation. The formulation, however, does not ensure a unique O-D solution because the formulation is not strictly convex in the O-D variables. Uniqueness of the solution is ensured if a target O-D is used (Sheffi, 1985; Yang et al., 1994; Fisk, 1989). These models require a complete set of link counts, a target O-D matrix and that the observed flows be in equilibrium. If traffic counts only approximate an equilibrium flow pattern then the O-D estimates obtained by the above methods are not realistic. The bias in the target O-D further deteriorates the quality of the O-D estimates.

O-D estimation through calibration of demand model: Most of the O-D estimation techniques belonging to this class assume that the trip making behavior is reasonably represented by a linear or non-linear model. Link count data is used to calibrate the demand model. The review of various O-D estimation methods through calibration of demand model is available in Nguyen (1984). These methods, require the knowledge of route choice proportions. The main criticism of this class of models is that by forcing the O-D to follow a demand relation the full information contained in link counts is not utilized.

Matrix estimation methods: These models use the link counts to estimate an O-D which is consistent with the observed flows and any prior information available about the O-D. Van Zuylen & Willumsen (1980) used entropy maximization for O-D estimation for uncongested networks. Fisk (1988, 1989) extended this work for congested networks. Multiobjective programming techniques were used by Gothe et al. (1989) to obtain the O-D. Other significant methods of this group (methods which do not use entropy formulation) include generalized least squares estimation by Bell (1991) and Cascetta (1984). Bierlair & Toint (1995) developed a methodology which uses generalized least squares and also allows the explicit use of data describing the structure of the O-D. Methods in this group require knowledge of route choice proportions.

PROPOSED METHODOLOGY

In the existing models of O-D estimation which utilize route choice proportions and link volumes, the route choice proportions are assumed to be exogenously determined and flow independent. These assumptions allow one to generate the route choice proportions required for the O-D estimation by using some proportional assignment technique like simple all-or-nothing assignment or the more general stochastic assignment techniques.

---

2Yang et al. (1992) suggested equilibrium assignment based bilevel optimization formulation which works with the information on partial link counts. Note that the O-D estimated by this method may not satisfy the flow constraints. Moreover the solution obtained to this complex optimization problem may correspond to a local optima.
However, critical assumption of flow independence is valid at best in the case of uncongested sparse networks (rural). In the case of dense networks with significant amount of congestion (urban centers), the above assumption’s validity is questionable. In such cases the use of route choice proportions generated by flow independent assignment techniques will lead to inaccurate O-D estimates. When the congestion is severe and only few link flows are used in the estimation of the O-D, the estimated O-D when assigned on to the network to obtain the flows on unobserved links will produce erroneous link flows. These observations necessitate the use of a flow dependent assignment procedure to endogenously generate the route choice proportions.

In this paper a multipath, fuzzy inference based, flow dependent assignment algorithm is proposed. The algorithm results in a bilevel optimization problem which is solved iteratively. A broad outline of the iterative assignment – O-D estimation model is provided in this section. Detailed discussion of the model is provided in a later section.

Outline of the Proposed Algorithm

In order to generate the route choice proportions \( p_{ija}^n \) in an iteration \( n \) for link \( a \), the O-D estimated in the previous iteration is assigned on to the network using the proposed assignment algorithm. These proportions are used to estimate the O-D in the present iteration. The process is repeated till the difference, \( E \), in the O-D matrices estimated in two successive iterations is not significant. The stepwise procedure of the assignment – O-D estimation is as follows:

**Step 0:** Initialize iteration number \( n = 0 \). Define the permissible change in O-D estimated in successive iterations, \( \varepsilon \). Obtain initial assignment proportions using a flow independent assignment technique (here STOCH algorithm is used (Dial, 1971)). Calculate \( T_{ij}^n \) (number of trips from \( i \) to \( j \) at iteration \( n \)) for all OD pairs.

**Step 1:** Assign the estimated O-D using the proposed assignment algorithm. Calculate the revised route choice proportions and go to Step 2.

**Step 2:** Increase the iteration counter by one. Estimate the O-D using observed flows. In this paper the following formulation is used to estimate the O-D:

\[
\text{Minimize } F(T) = \sum_{\gamma \in \text{OD}} T_{\gamma j}^n \ln(T_{\gamma j}^n)
\]

Subject to \( \sum_{\gamma \in \text{OD}} p_{\gamma a}^n T_{\gamma j}^n = V_a; \) \( V_a \) is the observed flow on link \( a \) \( T_{ij}^n \geq 0 \)

Calculate \( E \) using \( E = \sum_{\gamma \in \text{OD}} (T_{ij}^n - T_{ij}^{n-1})^2 \). If \( E \leq \varepsilon \) STOP,

Otherwise \( T_{ij}^n = (T_{ij}^n + T_{ij}^{n-1})/2 \) for all OD pairs and go to Step 1.

As shown in the above model, simple average of the O-D estimated in the present and previous iterations is used to obtain the revised route choice proportions to eliminate large fluctuations in \( p_{ija}^n \) that are estimated in successive iterations. As was discussed earlier, the route choice proportions are the key elements in the estimation of the O-D. It was also stated that the existing proportional assignment techniques to obtain the route choice
proportions does not work well when congestion on links of the network is significant. Thus a new assignment technique to obtain the proportions has been proposed here. In the next section this proposed assignment (route choice) model is described.

THE PROPOSED ASSIGNMENT TECHNIQUE

The proposed assignment technique is, as stated earlier, a multipath, fuzzy inference based, flow dependent model of the route choice behavior of the users of a network. In the next subsection the route choice behavior is briefly discussed as a justification of the proposed model. Later the proposed model is developed. Ideally, the description of the proposed algorithm should have been preceded by an introduction to the theory of fuzzy sets and fuzzy inference. This is not possible due to space restriction; for a good introduction the reader is referred to Zimmermann (1991).

Route Choice Behavior

A motorists choice for one route or the other is affected by many factors—some measurable and some non-measurable. Amongst these many factors, travel time, average travel speed, and type and condition of road are thought to be of primary importance to the motorist during the process of choosing a route. For example, one may choose a longer route if that route has lesser number of signalized intersections or the average travel speed is expected to be high. In general, the hypothesis which states that a motorist always chooses the shortest route (in distance) is not correct, not only because there are other factors which motivate the choice of a route but also because the shortest route itself changes with the expected flows (since one takes a longer time for the same distance). Assignment techniques like all-or-nothing which basically utilize the above hypothesis and do not dynamically modify the travel time on routes are thus improper models (especially for congested networks) of route choice behavior.

Multipath models on the other hand are better than all-or-nothing model because they acknowledge the fact that there are paths other than the shortest which a motorist considers in choosing a route. Logit and probit models fall in this class. Yet these models too, fail to accommodate the flow dependence on the attractiveness of a route.

Equilibrium assignment techniques which allow for flow dependence however suffer from certain other shortcomings when used as a part of the O-D estimation process (discussed earlier in Section 2). Another technique which also assumes flow dependence is the incremental capacity restraint method. This method, however, does not guarantee the convergence of the estimated flows to observed flows and therefore is unacceptable as a model to be used in the O-D estimation. One may refer to Sheffi (1985) for a detailed discussion on these and other multipath assignment models.

It is stated earlier that a motorist chooses his route not only based on time but he considers many other factors like condition of road, speed etc. Perceptions of the condition of a road and the travel speed can hardly be modeled deterministically or probabilistically. Perceived factors including the above are better characterized by fuzzy sets. Further, a motorist chooses that route which he/she infers to be the best under the above perceived conditions. Human inference procedures are better captured by the constructs of an approximate reasoning system like fuzzy inference. Thus, an appropriate model of route
choice behavior should be based on some approximate reasoning technique such as fuzzy inference technique.

Hence, the proposed model does not require path enumeration; considers more than one path in the choice set; incorporates the dependency of choice of route on expected flows; and uses an approximate reasoning procedure to more appropriately model the human inference process which precedes the decision of choosing a route.

Proposed Fuzzy Inference Based Assignment Model

The proposed multipath assignment technique, instead of considering all paths between an O-D pair considers only a set of paths which are termed as reasonable paths. Consider an O-D pair \((r, s)\) to be assigned on to a given network. Reasonable paths between \(r\) and \(s\) are those which are entirely constituted by links that take the traveler closer to the destination; i.e., each reasonable path from \(r\) to \(s\) includes only those links \(i \rightarrow j\) such that \(S(i) \geq S(j)\), where \(S(i), S(j)\) denote the travel time from nodes \(i\) and \(j\) to destination node \(s\) along the minimum path (from Dail’s STOCH algorithm, Dial, 1971).

According to the proposed assignment algorithm the amount of flow a path carries is a function of the weights assigned to all of the links \((LW)\) constituting the path. The weight is assigned to a link based on its attractiveness. The attractiveness of a link \((LA)\) is defined as the product of time attractiveness \((TA)\) and congestion attractiveness \((CA)\) of that link. The time attractiveness, \(TA(i \rightarrow j)\), of a link \(i \rightarrow j\) is a function (given later) of the extra travel time one incurs by choosing link \(i \rightarrow j\) over a link which is on the shortest path to \(s\) from \(i\) (this definition is adopted from Dial, 1971). For the purpose of calculating \(TA(i \rightarrow j)\) the travel time on the link \(i \rightarrow j(time_{ij})\) is assumed to be the travel time on the link under free flow conditions. The congestion attractiveness of a link is a function of the severity of congestion on that link. The severity of congestion is determined through a set of fuzzy rules and it is explained later. In the following, the calculation of \(TA, CA, LA\) (from \(TA\) and \(CA\)), \(LW\) and link flow \((LF)\) are discussed.

Time attractiveness, \(TA\): The time attractiveness \(TA(i \rightarrow j)\) for a link \(i \rightarrow j\) is

\[
TA(i \rightarrow j) = \exp(\theta LP(i \rightarrow j)),
\]

where, \(\theta\) is the slope parameter defining the sensitivity of \(TA\) to \(LP\); and \(LP(i \rightarrow j)\) is the extra travel time one incurs by choosing the path \(\{i \rightarrow j, \xi(j)\}\) instead of choosing \(\xi(i)\) (i.e., the time penalty for choosing \(i \rightarrow j\)). \(\xi(i)\) is the shortest path from \(i\) to destination node \(s\). \(LP(i \rightarrow j)\) needs to be calculated only for those links \(i \rightarrow j\) for which \(S(i) \geq S(j)\), and is given by

\[
LP(i \rightarrow j) = S(i) - S(j) - time_{ij}.
\]

Note that \(TA(i \rightarrow j)\) is flow independent and is obtained from static features of the network.

Congestion attractiveness, \(CA\): The congestion attractiveness, \(CA\) of a link is calculated based on the severity of congestion on that link. Fuzzy rules are used to determine the severity of the congestion. The following discussion illustrates the methodology used in the determination of \(CA\).
The fuzzy rules try to relate the length of the link and its level of congestion to the \( CA \) of that link. Note that the length of a congested link is taken as a surrogate measure of the duration for which a traveler faces the congestion. The length of the link and its congestion level are therefore premise variables of the rules and the congestion attractiveness the consequence of the rules. First the premise variables are presented. This is followed by a description of the consequence variables and finally the rules.

As shown in Figure 1(a), five linguistic classes, namely, very long (VL), long (L), average (A), short (S) and very short (VS), are defined for the premise variable, link length, LL. As shown in Figure 1(b), five linguistic classes, namely, very high congestion (VHC), high congestion (HC), average congestion (AC), less congestion (LC) and very less congestion (VLC), are defined (using link flow, \( LF \)) for the premise variable congestion.

For determining severity of congestion 25 variables are defined for consequences of rules. These are: \( w_1, w_2, \ldots, w_{25} \), where \( w_i \) represents various degrees of attractiveness from highly unattractive \( w_1 \) to highly attractive \( w_{25} \). In this paper \( w_1 \) through \( w_{25} \) are represented as triangular membership functions of fixed support and are symmetric about \( W_1 \) through \( W_{25} \). Note that \( p_{11}, p_{12}, \ldots, p_{15}, p_{21}, p_{22}, \ldots, p_{25} \) and \( w_1, w_2, \ldots, w_{25} \), are the parameters in all 25 rules. The following is an example of a typical rule.

**Rule 1:** If link is VL and congestion is VHC then link is highly unattractive \( (w_1) \)

For a given link \( i \rightarrow j \) (based on its length and the amount of flow it is carrying) one or more of the 25 rules are activated. The congestion attractiveness \( CA(i \rightarrow j) \) of the link is defined as the weighted average as follows.

\[
CA(i \rightarrow j) = \frac{\sum_{k=1}^{25} \alpha_k W_k}{\sum_{k=1}^{25} \alpha_k},
\]

where \( \alpha_k \) is the degree to which rule \( k \) is activated (\( \alpha_k = 0 \) means a rule is not activated and \( \alpha_k = 1 \) means a rule is fully activated). \( CA(i \rightarrow j) \) needs to be calculated for all \( i \rightarrow j \) for which \( S(i) \geq S(j) \). Note that \( CA(i \rightarrow j) \) is dependent on the flow.

**Link attractiveness, LA:** The link attractiveness is the product of \( TA \) and \( CA \) and is

\[
LA(i \rightarrow j) = TA(i \rightarrow j)CA(i \rightarrow j).
\]

![FIGURE 1. Membership functions (\( \mu \)) for (a) link length and (b) congestion.](image)
Link weights, \( LW \): The weight of a link \( i \to j \), \( LW(i \to j) \) is obtained from its attractiveness using the following equation (Dial, 1971).

\[
LW(i \to j) = LA(i \to j) \text{ if } j \text{ is } s \\
= LA(i \to j) \sum_{k \in D_j} LW(j \to k) \text{ otherwise,}
\]

where \( D_j \) is the set of links leaving node \( j \).

Link flow, \( LF^m \): The flow on link \( i \to j \), \( LF^m(i \to j) \) is calculated from its weight using (Dial, 1971)

\[
LF^m(i \to j) = OD_{r \to s}LW(i \to j)/ \sum_{k \in D_j} LW(i \to k) \text{ if } i \text{ is } r \\
= \sum_{k \in U_i} LF(I \to i)[LW(i \to j)/ \sum_{k \in D_i} LW(i \to k)] \text{ otherwise,}
\]

where \( U_i \) is the set of links entering node \( i \). The calculation of link flows should start from the origin node \( r \). Note that before calculating \( LF^m(i \to j) \) the flow on all the links coming to node \( i \) should be calculated.

**Summary of the Model**

The discussions in the preceding sections are presented here in the form of an algorithm.

**Step 0:** Define \( D_r, U_i \) for all the nodes \( i \). Set iteration index \( m=0 \). Initialize \( LF^m(i \to j) = 0 \). Define \( \delta \) as permissible flow deviation for convergence.

**Step 1:** Develop the skin tree from destination node \( s \), and determine \( S(i) \) for each node \( i \). Calculate \( TA(i \to j) \) for all links.

**Step 2:** Calculate, \( LA(i \to j) \), the link attractiveness.

**Step 3:** Calculate, \( LW(i \to j) \), the link weights.

**Step 4:** Calculate, \( LF^m(i \to j) \), the link flow.

**Step 5:** Calculate the flow difference index \( FDI \) using Eq. (7).

**Step 6:** If \( FDI \leq \delta \) STOP otherwise go to Step 2.

\[
FDI = \sum_{i \to j} |LF^m(i \to j) - LF^{m-1}(i \to j)|, \forall i \to j. \tag{7}
\]

In this section while presenting the procedure to calculate \( LA(i \to j) \), etc., it was assumed that the parameter \( \theta \) and parameters \( p_{11} \) through \( p_{25} \) and \( W_1 \) through \( W_{25} \) are known. However, for a given network and flow pattern these parameters need to be calculated. The procedure for obtaining these parameters is discussed next. It should be noted that the effect of the number of classes on \( FDI \) has not been studied here as the basic thrust of this paper is to propose the methodology and not a specific model. In this paper link
congestion along with travel times are considered to model the route choice but the methodology can incorporate other attributes which are significant in route choice.

**Calibration of the Assignment Algorithm**

The parameters of the proposed assignment algorithm are $\theta$ (required for calculating $TA$) and those which define the membership functions of the premise variables and consequence variables (required for calculating $CA$). In order to assign the O-D elements on to the network we need the values of the above parameters. For fixing the values of these parameters the assignment algorithm is to be calibrated and the parameters are to be so chosen that the resulting estimated flows are very close to the observed flows.

The most difficult part of the calibration (choosing the most appropriate values for the parameters) is the determination of the parameters pertaining to the fuzzy inference rules. In the past neural networks (back propagation algorithm) and genetic algorithms (GA) were used for the calibration of fuzzy rules (Kikuchi & Chakroborty, 1995 and Park & Kandel, 1994). The back propagation algorithm however, requires the consequences of the fuzzy rules to be measurable. As the consequences used in this study are non-measurable in nature, the back propagation algorithm cannot be used; hence GA are used. In the following a brief introduction to GA is provided. The reader may refer to Goldberg (1989) for a detailed discussion on GA.

**Introduction to GA**

GA are search techniques which work with a random population of points. Each point in GA represents a possible solution to the problem in question and a scheme is used for coding each point. Each coded point is called an individual or chromosome and consists of a list of genes, where each gene can assume finite number of values (1 or 0 in case of binary coded GA). GA work through an iterative procedure (each iteration is called a generation). GA derives their search power from clever combination of good points to obtain better points. In GA, first, each point in the population is evaluated to obtain their fitness (either by using an objective function or through some other procedure). After evaluating all the points in the population of a particular generation, a set of operators is used to generate new points. This new set of points constitute the population for the next generation.

A simple GA uses three basic operators to generate a new population of points; these are, reproduction, crossover and mutation. In the following a brief discussion on the operators is provided.

Reproduction is an operation through which individuals are copied into the mating pool based on its fitness. In the present paper stochastic reminder selection (SRS) operator is used for reproduction. If $f$ is the fitness of individual $k$ and $F$ is the average fitness of the population then by proportional selection then number of copies of individual $k$ in the next generation is $f/F$. In SRS operator, the number of copies of each individual in the population of next generation is first set equal to the integer part of $f/F$. In order to fill the remaining part of the population a biased roulette wheel is created with the strength of each individual equal to the decimal part of $f/F$ and a roulette wheel selection is done. The reproduction operator only emphasizes the best individuals, and does not give any search power to GA. The search power to GA comes predominantly from the crossover operator.

Crossover is the operation where two individuals (parents) are selected randomly from the mating pool (i.e., after reproduction) and the genetic information of two selected
individuals is exchanged at the selected sites to get two new individuals (children). In the present study a single point crossover (SPC) is used. In SPC a random number is generated between 1 and $l - 1$ where $l$ is the length of the chromosome, to select a crossover site. The genetic information to the right of the crossover site is exchanged to get two new individuals. In order to preserve some of the individuals, crossover is done only on a certain percentage of points in the mating pool called crossover probability. The SPC is illustrated with the following simple example constructed on two binary coded chromosomes. The vertical arrow represents the crossover site.

Parent one: 11100111  Child one: 11101010
Parent two: 01001010  Child two: 01000111

Mutation is done at the level of genes on the chromosomes obtained after crossover. Each gene of a selected chromosome is allowed to mutate to the other possible value (i.e., if the gene is 0, it will change to 1) with a certain probability known as mutation probability. The mutation operation is illustrated in the following example using the child one created in the above example. The vertical arrow is the mutation site.

Child one before mutation: 11101010
Child one after mutation: 10101010

**Calibration Procedure**

In the present work the fitness of an individual is calculated using the function:

$$f = \frac{1}{1 + \sum_{i,j \in L} (LF^m(i \rightarrow j) - LF^*(i \rightarrow j))^2};$$

where $LF^*(i \rightarrow j)$ is the observed flow on link $i \rightarrow j$ and $L$ is the set of links which are used for the calibration of the assignment algorithm. Note that in order to calculate the fitness of each individual one has to go through the entire procedure of assigning the O-D using the assignment algorithm with parameters as obtained from the chromosome. This is a costly operation and should be performed only a few times. Keeping this in mind a variation of the simple GA is used here. The variation used here is referred to in the literature as Micro GA (MGA). MGA can be thought of as simple GA operating with small population size, high mutation probability (to give more explorative power to MGA) and less number of generations (Krishnakumar, 1989). In order to preserve the best schema (best chromosomes), in the present study, two types of preselection are used at the reproduction stage; they are, replacing the worst string in the initial population by global best string found in the previous iteration of the O-D estimation process and replacing the worst string of the present generation by global best found so far in the O-D estimation process of the present iteration.

In the present MGA procedure the population size is 6, mutation probability is 0.1 and 0.125, crossover probability is 0.9 and maximum number of generations (M) are 10. Binary coding (see Goldberg, 1989) is used with string length 235 (10 bits for $\theta$ and for each of the premise variables of the fuzzy rules and 5 bits for each of the consequence variables of fuzzy rules); and $\beta$ is taken as 2. The calibration process proceeds as follows: first, initialize the population of MGA, then decode each chromosome to get the assignment parameters;
with the parameters obtained from a particular chromosome use the assignment algorithm and find the fitness of the string; with the fitness value perform GA operations; repeat the process for a predefined number of generations.

RESULTS

In order to illustrate the effectiveness of the proposed approach, results from two test cases are provided in this section.

Results from Case I

The network used for Case I is shown in Figure 2. In Figure 2 the travel time on all the links except the link represented by bold arrow is 2 units whereas the travel time on the link represented by bold arrow is 4 units.

The purpose of this case study is to point out the impact of link proportions on the O-D estimates. An attempt is made to show that the accuracy of the O-D estimates obtained from the existing procedures in this class of O-D estimation techniques are highly dependent on the flow pattern in the network. For example, if the flow pattern matches the underlying assumptions of an existing assignment algorithm then the estimates of the O-D elements obtained using link proportions generated by the same assignment algorithm are good, otherwise they are not. Further it is shown that the proposed assignment algorithm is versatile in that it can replicate various different flow patterns. Therefore, the O-D estimates obtained using the link proportions generated by the proposed algorithm are good for a variety of flow patterns.

The assignment algorithms used to illustrate the above claims are, (i) incremental capacity restraint assignment, (ii) Dial’s STOCH, and (iii) the proposed technique. In the CR technique the BPR relation \(\text{time}^{Va}_a = \text{time}^0_a [1 + 0.15(V_a/C_a)^4]\); where \(\text{time}^{Va}_a\) is the travel time on link \(a\) when the flow on link is \(V_a\), \(\text{time}^0_a\) is the travel time corresponding to free flow conditions and \(C_a\) is the capacity of the link \(a\) is used to update the travel times on the link. Ten increments are used to assign the O-D on to the network. In the STOCH algorithm the disutility coefficient associated with travel time is taken as unity for calculating the link attractiveness.

The study proceeds as follows:

1. A flow pattern is generated on the network for the true O-D (the \(T_{ij}\)’s for the true O-D are shown in parenthesis in Table 1). Three such flow patterns, namely Flow
Flow pattern I, Flow pattern II and Flow pattern III are generated using incremental capacity restraint assignment (CR) technique, Dial’s STOCH and the proposed fuzzy techniques respectively.

2. Each flow pattern is now assumed to be truly existing on the network. For each flow pattern the link proportions are obtained by using either the CR or STOCH or fuzzy technique. Each of the assignment techniques are calibrated for the given flow pattern by using the flow on link 2 → 5.

3. Using the link proportions obtained for a particular flow pattern and assignment technique, the O-D elements are estimated. The estimation procedure uses the flows on links 1 → 2 and 2 → 4.

The results obtained are shown in Table 1. It can be seen from the table that the estimates of the O-D elements obtained by using the proportions derived from CR or STOCH techniques are good only when the flow pattern (which is assumed to exist) on the network is also created by the same assignment technique. On the other hand, proposed technique consistently gives good O-D estimates irrespective of the flow pattern.

**Results from Case II**

Having shown the versatility of the proposed approach through earlier results, the proposed approach is applied to two larger networks to show that it can handle larger networks equally effectively. Figure 3 shows the first network used in this case. In the figure the following are the travel times on links: solid line — 1.1 units, dotted line — 1.2 units, dotted with center dash (for example link 6 → 10) — 0.8 unit, bold solid — 0.5 unit and bold dotted with center dash (for example link 7 → 11) — 0.9 unit. The flow pattern on the network is generated using STOCH algorithm; this pattern is assumed to exist on the network (it has been already shown that the efficiency of the proposed approach is not affected by any particular assumption related to the flow pattern. Here the choice of STOCH is truly arbitrary and is not important to the process of estimating the O-D elements). Twelve links (out of a total of 64 links in the network) namely 1 → 6, 6 → 10, 6 → 11, 7 → 10, 7 → 11, 7 → 12, 8 → 11, 8 → 12, 9 → 12, 10 → 15, 9 → 13 and 13 → 18 are used for calibrating the assignment algorithm. Flows on all the above links except links 6 → 10 and 10 → 15 are used in the estimation of O-D.

<table>
<thead>
<tr>
<th>Flow pattern</th>
<th>Method to obtain $p_{ij}^n$</th>
<th>Estimated O-D elements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T_{1,6}$ $T_{1,7}$ $T_{3,6}$ $T_{3,7}$</td>
<td></td>
</tr>
<tr>
<td>I (obtained by CR)</td>
<td>CR STOCH Fuzzy</td>
<td>1000 (1000) 1000 (1000) 500 (500) 500 (500)</td>
</tr>
<tr>
<td>II (obtained by STOCH)</td>
<td>CR STOCH Fuzzy</td>
<td>1000 (1000) 1000 (1000) 1000 (500) 1000 (500)</td>
</tr>
<tr>
<td>III (obtained by Fuzzy)</td>
<td>CR STOCH Fuzzy</td>
<td>1000 (1000) 1000 (1000) 500 (500) 500 (500)</td>
</tr>
</tbody>
</table>
The estimated O-D is given in Table 2. The numbers in the parenthesis in Table 2 correspond to the true O-D considered in this case. As can be seen from the table, the estimated O-D elements are close to their true values. The sum of absolute deviations \(\sum |\text{true O-D element} - \text{estimated O-D element}|\), for all OD pairs, is 309—a error of 2.06%.

Figure 4 shows the second network considered under Case II. In the network the link travel times are: dotted—1 unit, solid—1.2 units, vertical solid—0.8 unit, vertical dotted—0.9 unit, bold—1.3 units and bold and dotted—1.1 units. As earlier the flow pattern is generated using STOCH algorithm. Flows on 16 links (out of a total of 110 links)

### Table 2. Estimated O-D Elements for First Network (Case II)

<table>
<thead>
<tr>
<th>Origin node</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>167 (167)</td>
<td>221 (208)</td>
<td>258 (250)</td>
<td>286 (292)</td>
<td>342 (333)</td>
</tr>
<tr>
<td>2</td>
<td>206 (233)</td>
<td>302 (292)</td>
<td>349 (350)</td>
<td>390 (408)</td>
<td>487 (467)</td>
</tr>
<tr>
<td>3</td>
<td>206 (200)</td>
<td>241 (250)</td>
<td>270 (300)</td>
<td>325 (350)</td>
<td>450 (400)</td>
</tr>
<tr>
<td>4</td>
<td>278 (267)</td>
<td>328 (333)</td>
<td>410 (400)</td>
<td>480 (467)</td>
<td>517 (533)</td>
</tr>
<tr>
<td>5</td>
<td>144 (133)</td>
<td>165 (167)</td>
<td>199 (200)</td>
<td>226 (233)</td>
<td>266 (267)</td>
</tr>
</tbody>
</table>

### Table 3. Estimated O-D Elements for Second Network (Case II)

<table>
<thead>
<tr>
<th>Desti. node</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>211 (210)</td>
<td>289 (250)</td>
<td>234 (267)</td>
<td>327 (333)</td>
<td>305 (300)</td>
<td>493 (500)</td>
<td>828 (875)</td>
</tr>
<tr>
<td>27</td>
<td>179 (100)</td>
<td>90 (125)</td>
<td>406 (400)</td>
<td>658 (667)</td>
<td>800 (750)</td>
<td>659 (583)</td>
<td>571 (583)</td>
</tr>
<tr>
<td>28</td>
<td>102 (150)</td>
<td>252 (250)</td>
<td>1068 (1000)</td>
<td>500 (500)</td>
<td>481 (500)</td>
<td>170 (233)</td>
<td>302 (293)</td>
</tr>
<tr>
<td>29</td>
<td>378 (375)</td>
<td>666 (667)</td>
<td>642 (667)</td>
<td>175 (200)</td>
<td>245 (250)</td>
<td>372 (350)</td>
<td>588 (583)</td>
</tr>
</tbody>
</table>
in the network), namely, $3 \rightarrow 9$, $3 \rightarrow 10$, $5 \rightarrow 12$, $7 \rightarrow 13$, $10 \rightarrow 11$, $13 \rightarrow 19$, $14 \rightarrow 6$, $15 \rightarrow 21$, $18 \rightarrow 24$, $21 \rightarrow 26$, $22 \rightarrow 23$, $22 \rightarrow 26$, $23 \rightarrow 27$, $23 \rightarrow 28$, $24 \rightarrow 25$ and $25 \rightarrow 29$ are used to calibrate the assignment algorithm. All the above links except link $3 \rightarrow 9$ are used in the process of estimating the O-D.

The estimated O-D matrix is shown in Table 3. The numbers in the parenthesis of Table 3 correspond to the true O-D considered in this case. The estimated values are close to their
true values. The sum of absolute deviations is 739—an error of 3.1%. It should be understood that the parameters of the proposed assignment algorithm are calibrated using GA, which work probabilistically. However, the parameters do converge to approximately the same values after calibration. This fact can be seen from Figure 5. In this figure the two lines labeled Network 1 show how the sum of absolute deviations of the estimated O-D matrix for the first network converged to approximately the same value for two different runs of calibration process. The other two lines labeled Network 2 show the same for the second network. The purpose of this figure is to highlight that the O-D estimation process is not sensitive to the probabilistic nature of the calibration process.

CONCLUSION

The shortcomings of the various existing assignment techniques when used as a part of the O-D estimation process have been discussed. A new fuzzy inference based assignment approach to obtain the relative contributions of the O-D elements towards the link volumes is proposed. The proposed methodology for O-D estimation is tested on three test networks and the results are found to be promising. The results highlight that the proposed methodology is versatile as it can estimate the O-D elements well, for a variety of flow patterns. Further, the results show that the proposed approach holds promise for successful application to large networks with complex flow patterns.

REFERENCES


