Optimal Locations of License Plate Recognition to Enhance the Origin-Destination Matrix Estimation

Yu-Chiun CHIOU
Professor,
Institute of Traffic and Transportation
National Chiao Tung University
4F, 118 Sec.1, Chung Hsiao W. Rd., Taipei,
Taiwan 10012
Fax: +886-2- 2349-4953
E-mail: ycchiou@mail.nctu.edu.tw

Lawrence W. LAN
Chair Professor, Ta Hwa Institute of Technology; Emeritus Professor, National Chiao Tung University
No. 1 Tahwa Rd., Chuinglin, Hsinchu,
Taiwan 30740
Fax: +886-2- 2337-5755
E-mail: lawrencelan@thit.edu.tw

Chun-Ming TSENG
Ph.D. Candidate
Institute of Traffic and Transportation,
National Chiao Tung University,
4F, 118 Sec.1, Chung Hsiao W. Rd., Taipei,
Taiwan 10012
Fax: +886-2-2909-7032
E-mail: cm_tseng@mocom.gov.tw

Chih-Chao FAN
Postgraduate student
Department of Transportation Technology and Management
Feng Chia University
100, Wenhwa Rd. Seatwen, Taichung 40724
Taiwan
Email: smalleye106@yahoo.com.tw

Abstract: This paper develops a model to determine and to locate the optimal numbers of license plate recognition (LPR) to minimize error rates of O-D matrix estimation, percentages of LPR coverage (a proxy of installation cost), and percentages of recorded trips (a proxy of privacy invasion). A genetic algorithm is proposed to solve the combinatorial problem. To demonstrate the applicability of the proposed model and solving algorithm, two exemplified cases and a real-world case are investigated. The results have consistently showed that the optimal locations of LPR are at both ends and in the middle of the segments of a freeway corridor with heavy link traffics. If an extensive coverage of LPR is attempted, however, additional LPR may be placed at the segments with light link traffics so as to balance out the privacy invasion.

Key Words: license plate recognition, O-D matrix estimation, genetic algorithm

1. INTRODUCTION

Origin-destination (O-D) information is essential for traffic management and transportation planning. Conventionally, a large-scale roadside survey or household interview must be carried out to obtain O-D data, which is very costly and time-consuming. Taking advantages of traffic counts automatically collected from the link vehicular detectors, recently, many researchers have developed different algorithms to estimate the O-D matrices. However, O-D estimation basing on link traffic counts remains a challenging issue because the number of links with vehicular detectors is normally smaller than the number of O-D trips to be estimated. It is almost impossible to obtain a unique O-D matrix simply with the link traffic counts in most cases. Additional assumptions or exogenous information, such as route choice behaviors, a priori O-D matrix information, a sequence of observational periods of traffic counts data, may be required (see, for example, Bell, 1983, 1991; Yang et al., 1992, 1994;
Vardi, 1996; Lo et al., 1996; Hazelton, 2001). Different estimation models or techniques have been developed, such as generalized gravity model (Robillard, 1975), maximum entropy model (Fisk, 1989), generalized least squares (Bell, 1991), Bayesian estimation (Lo et al., 1996), artificial neural network (Yang et al., 1998), and bi-level programming (Yang, 1995).

The accuracy of O-D matrices estimated by these models or techniques depends heavily upon traffic patterns, network geometric layouts, and the rationales of assumptions. The estimation performance can hardly be guaranteed or enhanced without introducing further information. Today, license plate recognition (LPR) technique has become promising to provide additional information to enhance the O-D matrix estimation. Basically, the LPR system is an image-processing technology used to identify or track vehicles by their license plates for the purposes of security and traffic management including access control, traffic enforcement, and data collection. LPR can automatically record and check the license plate numbers of any passing vehicles and determine the partial trails of the vehicles. With the extra information, the accuracy of O-D matrix estimation might be remarkably enhanced. Obviously, if the LPR system is ubiquitous on all ramps of a freeway or at all intersections of a network, the O-D matrix is likely to be exactly estimated. However, an extensive installation of LPR may not be accepted by the road users due to the privacy invasion. More importantly, it is financially infeasible due to the government budget constraint. Therefore, it becomes an interesting and challenging topic to determine the appropriate number of LPR and their optimal locations so as to balance out the O-D estimation accuracy with acceptable privacy invasion and feasible installation cost.

Based on this, the present paper proposes a multi-objective mathematical programming model to determine the optimal locations of LPR, which minimizes three objectives: error rates of O-D matrix estimation, LPR installation cost, and privacy invasion. Genetic algorithm (GA) is employed to solve this combinatorial problem. A pseudoinverse technique is further proposed to estimate the O-D matrix based on the traffic information provided by the loop detectors and the LPR at some selected locations. These extra traffic equations are derived from a pairwise comparison of the recognized license plate numbers between any two arbitrary LPR. To investigate the applicability and effectiveness of our proposed model and solving algorithms, one small-scale problem and one large-scale problem are exemplified, and one real-world case study are conducted.

The remaining parts of this paper are organized as follows. Section 2 formulates the model and gives a seminal example to demonstrate how the traffic equations can be derived from a pairwise comparison of license plate numbers recognized by different LPR. Section 3 presents the cases of two different-scale computational examples to investigate the effectiveness of proposed model. Section 4 further demonstrates the applicability of the proposed model on the real-world freeway context. Finally, the concluding remarks and suggestions for future studies are addressed in Section 5.

2. MODEL

2.1 Problem Formulation

A linear freeway corridor denoted as $G(N, A)$ with an O-D matrix $T$, where $N$ is a set of interchanges and $A$ is a set of segments (between two interchanges). Each segment is a
candidate location to install exactly one set of LPR. Assume that each of the freeway segments has already been embedded with at least one set of loop detectors by which the basic information of link traffic counts is provided. The LPR problem formulation is to minimize the following three objectives: error rate, installation cost and privacy invasion:

\[
\begin{align*}
\text{Min} & \quad E = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{|L_{ij} - \hat{L}_{ij}|}{t_{ij}} \\
\text{Min} & \quad C = \sum_{k=1}^{A} y_k c \\
\text{Min} & \quad P = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \delta_{ij} t_{ij}
\end{align*}
\]

Subject to:

\[
\sum_{i=1}^{N} \sum_{j=k+1}^{N} t_{ij} = x_k , \text{ for all } k, i, j. 
\]

\[
|A| \delta_{ij} \geq \sum_{k=1}^{N} y_k , \text{ for all } k, i, j. 
\]

\[
\delta_{ij} \leq \sum_{k=1}^{N} y_k , \text{ for all } k, i, j. 
\]

\[
\hat{L}_{ij} = f(y_1, y_2, ..., y_k, ..., y_{|A|}) , \text{ for all } i, j. 
\]

\[
y_k = \{1, 0\} , \text{ for all } k. 
\]

where \(y_k\) is a binary decision variable: 1 denotes one set of LPR installed at segment \(k\); 0, otherwise. \(x_k\) is the traffic counts from the loop detectors at segment \(k\). \(t_{ij}\) represents the real traffic from interchange \(i\) (origin) to interchange \(j\) (destination), which is the \((i, j)^{th}\) element of the O-D matrix \(T\). \(\hat{L}_{ij}\) is the decision variable representing the estimated traffic from interchange \(i\) to interchange \(j\) for all \(i, j=1, 2, ..., |N|\) and \(j > i\). \(\delta_{ij}\) is a dummy variable: 1 represents the O-D pair from \(i\) to \(j\) will be pictured and recorded by at least one LPR; 0, otherwise. \(|N|\) is the number of interchanges. \(|A|\) is the number of segments. \(c\) denotes the installation cost of one set of LPR. \(E\) denotes the error rate for estimating the O-D matrix. \(C\) denotes total installation cost of LPR. \(P\) denotes total number of trips pictured and recorded by LPR, which is a proxy of total amount of privacy being invaded.

A simple additive weighted method is employed to solve the above multi-objective programming problem, in which values of the three objectives are normalized into a range of \([0, 1]\).

\[
\text{Min } Z = \alpha \left( \frac{E - E_{min}}{E_{max} - E_{min}} \right) + \beta \left( \frac{C - C_{min}}{C_{max} - C_{min}} \right) + \gamma \left( \frac{P - P_{min}}{P_{max} - P_{min}} \right)
\]

where \(E_{max}\) is the maximal error rate determined by solving O-D matrix based on link traffic counts alone. \(E_{min}\) is the minimal error rate. If installing LPR at every segment of the freeway, then \(E_{min}=0\). \(C_{max}\) is the maximal LPR installation cost, so \(C_{max} = c|A|\). \(C_{min}\) is the minimal LPR installation cost. If none of LPR are installed, then \(C_{min}=0\). \(P_{max}\) is the maximal amount...
of privacy being invaded, i.e., \( P_{max} = \sum_{i=1}^{[Y]} \sum_{j=1}^{[Y]} t_{ij} \), \( P_{min} \) is the minimal amount of privacy being invaded. If none of LPR are installed, then \( P_{min} = 0 \). \( \alpha, \beta, \) and \( \gamma \) are the corresponding weights of three objectives, respectively.

Since the installation costs of LPR at various segments are the same, the second component of Eq. (9) can be interpreted as the percentage of LPR coverage. Likewise, the third component of Eq. (9) can also be interpreted as the percentage of recorded traffic. Eq. (4) is the traffic equations derived from loop detectors. Eq. (5) and Eq. (6) indicate whether there is at least one set of LPR being installed between interchanges \( i \) and \( j \). Eq. (7) represents the relationship of O-D matrix estimation based on traffic equations derived from license plate comparisons among LPR. It is difficult to derive a generalized explicit form as a function of \( y_k \), thus \( f(\cdot) \) is used to implicitly represent this relationship, which is explained in section 2.2.

In most cases, the number of traffic equations is less than the number of O-D trips to be estimated; thus, this paper employs pseudoinverse technique, also known as generalized inverse or Moore-Penrose inverse technique, to estimate O-D trips based on these traffic equations. The pseudoinverse technique is commonly used to solve a minimal-length solution (closest to the origin) of the underdetermined linear equations problem. Assume that a total of \( m \) traffic equations can be obtained from link traffic counts and license plate comparisons as follows:

\[
\begin{align*}
X &= F\hat{T} \\
\end{align*}
\]

where \( X \) represents a vector of link traffic counts and license plates comparisons; \( \hat{T} \) denotes a vector of \( p \) estimated O-D trips; \( F \), a \( m \times p \) matrix, represents the coefficients of these traffic equations. Since the number of traffic equations is less than the number of estimated O-D trips, \( F \) is not a square matrix. The estimated O-D trips can be solved by the following equation:

\[
\hat{T} = F'(FF')^{-1}X
\]

where \( F' \) denote a transpose of matrix \( F \), \( F'(FF')^{-1} \) is a pseudoinverse of matrix \( F \).

2.2 A Seminal Example

To demonstrate how the traffic equations can be derived from a pairwise comparison of license plate numbers recognized by different LPR, this paper considers a stretch of eastbound freeway with six interchanges and five loop detectors (Figure 1). Each interchange can be the origin or the destination of the trips. In the case that two sets of LPR are installed at the 2nd and 4th segments, a total of 15 unknown \( \hat{i}_{ij} \) have to be determined.
Five traffic equations can be derived from the link traffic counts of loop detectors by the equation of $\sum_{i=1}^{k} \sum_{j=k+1}^{n} t_{ij} = \sum_{j=k+1}^{n} t_{j}$; namely,

\[
\begin{align*}
&x_1 = t_{12} + t_{13} + t_{14} + t_{15} + t_{16} \\
&x_2 = t_{13} + t_{14} + t_{15} + t_{16} + t_{23} + t_{24} + t_{25} + t_{26} \\
&x_3 = t_{14} + t_{15} + t_{16} + t_{24} + t_{25} + t_{26} + t_{34} + t_{35} + t_{36} \\
&x_4 = t_{15} + t_{16} + t_{25} + t_{26} + t_{35} + t_{36} + t_{45} + t_{46} \\
&x_5 = t_{16} + t_{26} + t_{36} + t_{46} + t_{56}
\end{align*}
\]

Three additional traffic equations can be derived from these two LPR by comparing the license plate numbers recognized:

\[
\begin{align*}
&\|l_2 \cap l_4\| = t_{15} + t_{16} + t_{25} + t_{26} \\
&\|l_2 - l_2 \cap l_4\| = t_{13} + t_{14} + t_{23} + t_{24} \\
&\|l_4 - l_2 \cap l_4\| = t_{35} + t_{36} + t_{45} + t_{46}
\end{align*}
\]

where, $l_k$ represents a set of license plate numbers successfully recognized by the LPR located at the $k^{th}$ segment. $\cap$ is the intersection operator—a set of matching license plates recognized by these two LPR. For instance, $\|l_2 \cap l_4\|$ in Eq. (17) represents the number of matching license plates recognized by the LPR at segment 2 and 4.

If $x_1 = 50$, $x_2 = 120$, $x_3 = 180$, $x_4 = 190$, $x_5 = 150$, $\|l_2 \cap l_4\| = 60$, $\|l_2 - l_2 \cap l_4\| = 60$, $\|l_4 - l_2 \cap l_4\| = 140$, then the estimated O-D trips can be determined by using the pseudoinverse technique and the error rate of O-D matrix estimation is 23.93%. If an additional LPR is installed at the 3rd segment (Figure 2), six traffic equations can be further derived from the dual comparisons of license plates recognized by the paired LPR:

\[
\begin{align*}
&x_2 = t_{12} + t_{13} + t_{14} + t_{15} + t_{16} + t_{23} + t_{24} + t_{25} + t_{26} \\
&x_3 = t_{13} + t_{14} + t_{15} + t_{16} + t_{23} + t_{24} + t_{25} + t_{26} + t_{34} + t_{35} + t_{36} \\
&x_4 = t_{14} + t_{15} + t_{16} + t_{24} + t_{25} + t_{26} + t_{34} + t_{35} + t_{36} + t_{45} + t_{46} \\
&x_5 = t_{15} + t_{16} + t_{25} + t_{26} + t_{35} + t_{36} + t_{45} + t_{46} + t_{56}
\end{align*}
\]
\[ |l_2 \cap l_4| = t_{14} + t_{15} + t_{16} + t_{24} + t_{25} + t_{26} \]  \hspace{1cm} (20)  
\[ |l_2 - l_2 \cap l_3| = t_{13} + t_{23} \]  \hspace{1cm} (21)  
\[ |l_3 - l_2 \cap l_3| = t_{34} + t_{35} + t_{36} \]  \hspace{1cm} (22)  
\[ |l_3 \cap l_4| = t_{15} + t_{16} + t_{25} + t_{26} + t_{35} + t_{36} \]  \hspace{1cm} (23)  
\[ |l_3 - l_3 \cap l_4| = t_{14} + t_{24} + t_{34} \]  \hspace{1cm} (24)  
\[ |l_4 - l_3 \cap l_4| = t_{45} + t_{46} \]  \hspace{1cm} (25)  

where Eqs. (20)-(22) are derived from the comparisons of license plates recognized by the LPR at the 2\textsuperscript{nd} and 3\textsuperscript{rd} segments. Eqs. (23)-(25) are derived from the comparisons of LPR at the 3\textsuperscript{rd} and 4\textsuperscript{th} segments. In addition, another seven traffic equations can also be derived from the triple comparisons of license plates recognized by these three LPR:

\[ |l_2 \cap l_3 \cap l_4| = |l_2 \cap l_4| \]  \hspace{1cm} (26)  
\[ |l_2 \cap l_3 - l_2 \cap l_3 \cap l_4| = t_{14} + t_{24} \]  \hspace{1cm} (27)  
\[ |l_2 \cap l_4 - l_2 \cap l_3 \cap l_4| = 0 \]  \hspace{1cm} (28)  
\[ |l_3 \cap l_4 - l_2 \cap l_3 \cap l_4| = t_{35} + t_{36} \]  \hspace{1cm} (29)  
\[ |l_2 - l_2 \cap l_3 \cap l_4| = |l_2 - l_2 \cap l_4| \]  \hspace{1cm} (30)  
\[ |l_3 - l_2 \cap l_3 \cap l_4| = t_{14} + t_{24} + t_{34} + t_{35} + t_{36} \]  \hspace{1cm} (31)  
\[ |l_4 - l_2 \cap l_3 \cap l_4| = |l_4 - l_2 \cap l_4| \]  \hspace{1cm} (32)  

Note that Eq. (27) can be derived from Eqs. (18) and (21); Eq. (29) can be derived from Eqs. (19) and (25); Eq. (31) can be derived from Eqs. (18), (21) and (22). Namely, Eqs. (26)-(32) are in effect duplicated by the equations derived from pairwise comparisons. Thus, it is sufficient that one considers only the information provided by LPR pairwise comparisons.

If \[ |l_2 \cap l_1| = 90, \quad |l_2 - l_2 \cap l_1| = 30, \quad |l_3 - l_2 \cap l_3| = 90, \quad |l_3 \cap l_4| = 120, \quad |l_3 - l_3 \cap l_4| = 60, \quad |l_4 - l_3 \cap l_4| = 80, \] then the O-D matrix can be estimated with an error rate of 12.32%. This example demonstrates that the error rate can be reduced by 11.61% if an additional LPR is installed at segment 3.

2.3 Solution Algorithm

Due to the combinatorial nature of the problem, GA is suitable to solve the problem with the following eight steps:

Step 0: Initialization. Generate an initial population with \( M \) chromosomes. Each chromosome has \( |A| \) genes and each gene randomly takes an integer of \{1, 0\}. The value of gene is used to indicate whether an LPR is installed at the corresponding segment or not. Set the values of related parameters including population size, crossover rate, mutation rate, matured rate (\( \eta \)), and weights of three objectives (\( \alpha, \beta \) and \( \gamma \)).

Step 1: Estimation of O-D trips. Estimate the O-D trips based on the traffic equations derived from link traffic counts by loop detectors and license plate comparisons by LPR according to the corresponding chromosome.

Step 2: Evaluation. Calculate the fitness values of all chromosomes.
Step 3: Selection. The roulette wheel method is adopted.
Step 4: Crossover. The two-point crossover method is adopted.
Step 5: Mutation. The gene mutation method is adopted.
Step 6: Repeat Step 1 and Step 2.
Step 7: Testing of stop condition. If mature rate $\geq \eta$, then stop. The incumbent location is the optimal selection. Otherwise, go to Step 3.

3. EXEMPLIFIED PROBLEMS

To investigate the performance of the proposed model and the solving algorithm, two exemplified cases, one small-scale problem and one large-scale problem, are tested. In addition to use the proposed solving algorithm, the small-scale problem is also solved by total enumeration method to obtain the true optimal results, which are used to verify the effectiveness of the proposed solving algorithm.

3.1 Small-scale Problem

The small-scale problem considers a freeway of 8 interchanges (7 segments) with a randomly generated O-D matrix (Table 1). There are a total of 28 O-D trips, which are used to compute the link traffics of loop detectors and recognized traffics of LPR and also used to compute the error rate of the O-D matrix estimation. The parameter settings of GA are as follows: length of chromosomes=7, population size=50, crossover rate=0.8, mutation rate=0.01, and mature rate =0.8.

<table>
<thead>
<tr>
<th>$i$</th>
<th>$j$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>32</td>
<td>61</td>
<td>44</td>
<td>52</td>
<td>46</td>
<td>37</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>15</td>
<td>39</td>
<td>55</td>
<td>47</td>
<td>62</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>23</td>
<td>54</td>
<td>66</td>
<td>21</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>32</td>
<td>51</td>
<td>47</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>33</td>
<td>46</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>52</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>

To understand the marginal contribution of additional LPR, the O-D matrix estimation error rate is presented in the following with various preset numbers of LPR ($i.e. \sum_{k=1}^{d} y_k = n, n$ is the preset number of LPR). Solving the small-scale problem by both GA and total enumeration method, the results are presented in Table 2. Note that GA takes only one or two generations to converge and it yields the same optimal results as those solved by total enumeration method under various preset numbers of LPR. Without installation of any LPR, the error rate will go as high as 47.96%; however, the error rates are gradually reduced to 36.65%, 27.49%, 19.24%, 11.07%, and 0.00% as the numbers of LPR increase to 2, 3, 4, 5, and 6, respectively. In other words, the marginal contribution of additional LPR in reducing the error rates of O-D estimation ranges from 8.17% to 11.31%.
Table 2 Minimal $E$ under various numbers of LPR

<table>
<thead>
<tr>
<th>Method</th>
<th>Number of LPR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>GA</td>
<td>47.96%</td>
</tr>
<tr>
<td>Total enumeration method</td>
<td>47.96%</td>
</tr>
<tr>
<td>Marginal contribution (reduction of $E$) for additional LPR</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 3 further depicts the error rates of every enumeration under various numbers of LPR. It can be found that even under the same number of LPR, different locations of LPR have exhibited remarkably different error rates. Taking $n=3$ for instance, the largest error rate has reached 40.25%, much larger than the minimal one (27.49%), suggesting the importance of proper location of the LPR.

The optimal locations for various numbers of LPR are depicted in Figure 4. Note that the optimal locations of LPR are not installed at the segment with high link traffic; instead, they should be located at both ends of the corridor segments wherein link traffics are heavy (segments 3) so as to capture the longer-distance O-D trips.
3.2 Large-scale Problem

The large-scale problem considers a freeway of 42 interchanges (41 segments) with randomly generated O-D matrices, in which link traffic distribution is shown in Figure 5. The corresponding weights of three objectives in Eq. (9) are set as $\alpha = 0.5$, $\beta = 0.25$, and $\gamma = 0.25$. The parameter settings of GAs remain the same as the small-scale problem. The optimal locations corresponding to various numbers of LPR are reported in Table 3.

According to Figure 5 and Table 3, some interesting findings can be outlined. In case of low LPR coverage, the first best choice is to locate the LPR at both ends of high link traffic segments (e.g., 4, 15, 27, 40) so as to capture the partial trails of most vehicles. The second best choice is to locate the LPR at the middle of high link traffic segments (e.g., 6, 10, 15, 31). However, if a high LPR coverage is desired, the additional LPR should be located at the segments with light link traffic so as to avoid over-invading the privacy of road users. As a consequence, the distribution of optimal locations of LPR in high coverage cases is somewhat similar to the inverse curves of link traffic distributions.
which allows 49.09% error rate with 24.39% LPR coverage and 27.19% recorded trips.

Note: * represents a set of LPR installed.

### 4. CASE STUDY

To further demonstrate the applicability of the proposed model, a case study on Taiwan No.1 Freeway is conducted. This freeway is 372.7 kilometers in length with a total of 42 interchanges from Keelung (the very north end) to Kaohsiung (the very south end). The most updated O-D matrix of this freeway was surveyed by Institute of Transportation, Ministry of Transportation and Communications in 1995. The corresponding link traffic data at all segments are displayed in Figure 6. The corresponding weights of three objectives in Eq. (9) are set as $\alpha=0.5$, $\beta=0.25$, and $\gamma=0.25$. The parameter settings of GAs remain the same as the aforementioned large-scale problem. The results show that it requires installing 10 LPR, which allows 49.09% error rate with 24.39% LPR coverage and 27.19% recorded trips.

![Figure 6 Distribution of link traffic on Taiwan No.1 Freeway](source)

The optimal locations for various numbers of LPR are detailed in Table 4. As shown in Figure 6, an obvious peak traffic can be identified between segments 7 and 15, thus the first best choice for locating the LPR are at the segments 7 and 15, and then the middle of the peak.
(segment 10). As the number of LPR increases, it is likely to install the LPR at the segments with light link traffic to account for the privacy invasion factor.

Table 4 Optimal locations of LPR for Taiwan No.1 Freeway under various numbers of LPR

| No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 |
|-----|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
|     |   |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |

Note: * represents a set of LPR installed.

To look into the sensitivity of the corresponding weights of three objectives in Eq. (9), this paper varies $\alpha$ from 0.50 to 0.95 and sets $\beta = \gamma$. The optimal numbers of LPR associated with different values of $\alpha$ are presented in Table 5. Notice that as $\alpha$ gets larger the error rate decreases but the percentages of LPR coverage and recorded trips increase. When $\alpha$ is smaller than 0.5, the error rate is relatively high and the percentages of LPR coverage and recorded trips are relatively low, and the suggested numbers of LPR are less than 10. Also note that when $\alpha=0.95$ the error rate can be lowered to 9.16%, but the percentages of LPR coverage and recorded trips would reach as high as 70.43% and 75.43%, respectively.

Table 5 Optimal numbers of LPR associated with various values of $\alpha$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>Number of LPR</th>
<th>$E$</th>
<th>$C$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>10</td>
<td>49.09%</td>
<td>24.39%</td>
<td>27.19%</td>
</tr>
<tr>
<td>0.55</td>
<td>11</td>
<td>55.33%</td>
<td>26.83%</td>
<td>28.26%</td>
</tr>
<tr>
<td>0.60</td>
<td>11</td>
<td>43.08%</td>
<td>26.83%</td>
<td>28.83%</td>
</tr>
<tr>
<td>0.65</td>
<td>12</td>
<td>42.79%</td>
<td>29.27%</td>
<td>30.91%</td>
</tr>
<tr>
<td>0.70</td>
<td>15</td>
<td>38.48%</td>
<td>36.59%</td>
<td>38.91%</td>
</tr>
<tr>
<td>0.75</td>
<td>16</td>
<td>34.34%</td>
<td>39.02%</td>
<td>41.46%</td>
</tr>
<tr>
<td>0.80</td>
<td>18</td>
<td>31.99%</td>
<td>43.90%</td>
<td>45.03%</td>
</tr>
<tr>
<td>0.85</td>
<td>19</td>
<td>28.09%</td>
<td>46.34%</td>
<td>47.89%</td>
</tr>
<tr>
<td>0.90</td>
<td>24</td>
<td>19.17%</td>
<td>58.54%</td>
<td>62.66%</td>
</tr>
<tr>
<td>0.95</td>
<td>29</td>
<td>9.16%</td>
<td>70.73%</td>
<td>75.43%</td>
</tr>
</tbody>
</table>

5. CONCLUDING REMARKS

License plate recognition (LPR) is an innovative technique to automatically record and match the license plate numbers of passing vehicles so as to determine the partial trails of the vehicles. It adds into the conventional traffic counts with extra information to enhance the accuracy of O-D matrix estimation. The present paper has proposed a model to locate the optimal numbers of LPR to minimize three objectives: error rates of O-D matrix estimation,
percentages of LPR coverage (a proxy of installation cost), and percentages of recorded trips (a proxy of privacy invasion). Due to the combinatorial nature, the paper proposed to use GA to solve the problem. According to the exemplified small-scale and large-scale problems and a field case study, consistent results were found: the optimal locations of LPR should be at both ends and middle of the segments with heavy link traffic if a small number of LPR is to be installed. In contrast, as the number of LPR increases, it is likely to install the LPR at the segments with light link traffic to balance out the privacy invasion.

Some directions for future studies can be identified. Firstly, this paper employed a rather simple method—pseudoinverse technique to investigate the marginal contribution of additional LPR in estimating the O-D matrix. Future study can incorporate other logical assumptions, such as route choice behaviors or user equilibrium, into more complex techniques, such as generalized least squares or bi-level programming, so as to further improve the accuracy of O-D matrix estimation. Secondly, it deserves further elaborating the proposed model to determine the optimal locations of LPR to accommodate the real-world dynamic O-D matrices estimation. Last but not least, the proposed model can also be extended to a complicated network context.

ACKNOWLEDGEMENTS

This manuscript is moderately revised from an early work (Chiou et al. 2006) presented at the 11th International Conference of Hong Kong Society for Transportation Studies. The study was partially sponsored by the National Science Council, ROC (NSC 96-2628-E-009-171-MY3).

REFERENCES