A Fast Accurate LP Approach for Traffic Matrix Estimation

S. Eum, J. Murphy, and R. J. Harris

aCentre for Advanced Technology in Telecommunications (CATT Centre) School of Electrical and Computer Engineering, RMIT University, Melbourne GPO 2476V Victoria 3001, Australia Email: {suyong, murphy}@catt.rmit.edu.au

bInstitute of Information Sciences and Technology Massey University, Palmerston North Private Bag 11 222, New Zealand Email: R.Harris@massey.ac.nz

Abstract: There has been much interest recently in “Internet Tomography” - using inference methods to infer the size of the traffic flows in the traffic matrix because direct measurement of these flows has many technical difficulties. Previously, it was thought that Linear Programming (LP) methods produced poor results. We show, however, that using the interior-point method for Linear Programming, rather than the more traditional Simplex Method, we obtain very good results that are as fast and as accurate as the method of “Tomogravity” developed by AT&T.

Keywords: Traffic Matrix, Linear Programming, Interior Point Method, Tomogravity

1. INTRODUCTION

Traffic matrix estimation methods, known as “Internet Tomography” [1], rely upon solving systems of equations that are highly under-constrained. The number of unknown variables, which is the number of origin and destination pairs in a network, increases in proportion to the square of the number of nodes while the number of constraints, which is the number of links in a network, increases linearly. Therefore, as the size of the traffic matrix increases, the problem becomes increasingly under-constrained.

In [2], a deterministic LP [3] approach was compared to other traffic matrix estimation techniques. Interestingly, they concluded that this LP method generated such a high error, it made the method impractical. They reported that the LP method assigned zero values to many elements of the traffic matrix so it over-compensated by assigning very large values to the other elements of the traffic matrix. We observed this as well when the LP problem was solved by the Simplex Method. However, when we used the interior point method to solve it, we had much better results. The interior point method cuts across the interior of the feasible area to reach an optimum solution. Therefore, the method overcomes the restriction of the Simplex method which chooses an optimum solution from the corner points of the feasible space. This is possible because the optimal solution to this problem generally lies on a plane rather than at a single basic feasible solution (corner point of the convex hull).

There has recently been a promising traffic matrix estimation method published in the literature called Tomogravity [4]. Tomogravity is based on a combination of the well-known gravity model and the method of least squares. Importantly, the accuracy of the Tomogravity method was verified using a real traffic matrix and is being used for various network management tasks.

A comparison study between the interior point method and the Tomogravity method was conducted to test the accuracy of the interior point method. The authors developed a method to imitate a back-bone traffic matrix because there were not any real traffic matrices available. The simulation results showed that the accuracy of the interior point method was comparable to that of the Tomogravity or even better in some cases.
The rest of this paper is organized as follows: In Section 2, the problem of traffic matrix estimation using the LP and Tomogravity methods is illustrated. Implementations of the LP methods and the Tomogravity method are described in Section 3. In Section 4, a synthetic traffic matrix is proposed for simulating a back-bone network traffic matrix and an experiment is described that enables some comparisons to be made of the two approaches using two different topologies. Lastly, some results and conclusions are presented in Section 5.

2. ILLUSTRATION OF LP AND TOMOGRAVITY METHODS

In Fig. 1, the three node network has two links with three flows. These three flows need to be estimated from measurements of the two link loads which are 12 and 16 respectively. The sum of flows 1 and 2 is equal to the measured link loads which is 12 and the sum of flow 2 and 3 are the same as the measured link loads 16. The two constraints are illustrated in Fig. 2 based on Equations (1) and (2) respectively. This is an under-constrained problem because the number of unknown variables is more than the number of constraints.

\[
X_1 + X_2 = 12 \tag{1}
\]

\[
X_2 + X_3 = 16 \tag{2}
\]

Therefore, the problem defines a solution plane rather than giving a unique point for the solution. In Fig. 2, the line AB represents the solution which satisfies both constraints. Whatever technique is used for traffic matrix estimation, a solution from the method should lie on the line AB to satisfy the inter link measurement constraints. In the LP method case, if a linear objective function is selected that includes \((X_1, X_2, X_3)\) but has a different gradient from the line AB, then this objective function will give a unique answer which would be at either point A or point B of the Fig. 2. However, if the gradient of the objective function is parallel to that of the solution line AB, then an infinite range of solutions is possible. Note, however, that the Simplex Method will give a unique answer of A or B when applied to this situation because it must terminate at an extreme point. However, the interior point method produces a solution located between A and B. For the Tomogravity approach, the initial solution D is obtained using the Gravity model and then a line is drawn perpendicular to the solution line AB from the initial solution D. Lastly, the point C, which is closest to the initial solution D, is the final solution obtained by the Tomogravity method.
3. LINEAR PROGRAMMING AND TOMOGRAVITY

3.1. LP Problem Modelling

Goldschmidt et al [3] have formulated the traffic matrix estimation problem based on a network-flow approach. An element of a traffic matrix is called a flow or a traffic demand which represents the amount of traffic from a given node to another node. Suppose that there is a network where \( B \) is the vector of link loads, \( A \) is a routing matrix, and \( X \) is the vector of the unknown traffic demands. Then the network system can be described by the vector equation: \( AX = B \). In Internet Protocol (IP) networks, these prior conditions can be achieved relatively easily. For instance, the routes can be obtained by implementing Dijkstra algorithm to find a shortest path between an origin and destination node and then they can be coded into the routing matrix \( A \); also, measured link loads are available from SNMP data in typical IP networks. Therefore, the LP model can be constructed with the following objective function and constraints:

\[
\text{Maximize } \sum_{1 \leq j \leq |X|} w_j X_j \tag{3}
\]

subject to

\[
\sum_{j=1}^{c} A_{\ell j} X_j \leq B_{\ell} \quad (\ell = 1, \ldots, m) \tag{4}
\]

\[
\sum_{j=1}^{n} M_{ij} = R_i \quad (R_i : \text{row sum}), \quad \sum_{i=1}^{n} M_{ij} = C_j \quad (C_j : \text{column sum}) \tag{5}
\]

where \( w_j \) is a weight for OD pair \( X_j \). In [3], the authors suggested setting the weights \( w_j \) to the length of a path, ie using the hop count. \( m \) is the number of links, and \( c \) is the number of flows on a link. Equation (4) states that the sum of the flows on a link should be less than or equal to the measured link loads.

In Equation (5), \( M \) is the \( n \times n \) traffic matrix (\( n \) nodes network) and \( M_{ij} \) shows an element of the matrix. A row total \( R_i \) of the traffic matrix represents the amount of traffic entering the network through node \( i \) and a column total \( C_j \) gives the amount of traffic leaving a network through node \( j \). These row and column totals can be measured at the edge links using SNMP measurement data.

3.1.1. Simplex Method

The Simplex Method is based on the fact that an optimal LP solution can always be associated with an extreme point of the feasible region. Therefore, the Simplex Method moves around the boundary of the feasible region to find an adjacent extreme point which improves the objective function. As a result, any possible solutions are ignored if they are not associated with an extreme point. In [2], the Simplex algorithm seems to be used because the authors claimed that the result from the LP method includes a lot of zero values which means the solution is related to an extreme point (All non-basic variables would be zero in such a solution.). We also obtained the similar result with [2] when the simplex method was used to solve the LP problem.

3.1.2. Interior Point Method

While the Simplex Method follows a path of adjacent extreme points along the edge of the feasible region, the Interior Point Method cuts across the interior of the feasible area to reach an optimal solution. The concept of the Interior Point Method has existed for a long time.
based on the “Affine Scaling Method of Dikin”, “the Logarithmic Barrier Method of Frisch”, and “the Center Method of Huard”. However, after Karmarkar [6] introduced his method called the Projective method, in 1984, the Interior Point Method (IPM) regained attention due to the fact that the number of iterations depends only to a small extent upon the problem size [5].

The IPM involves a barrier method which is used to set a numerical boundary in the problem so that it forbids any variable from being equal to this boundary. This enables the IPM to produce strictly positive solutions. When there is an objective function $c^T x$ to be minimized with $x > 0$ constraint, the problem can be written as follows.

$$f(x; \mu) = c^T x - \mu \sum_{j=1}^{n} \log x_j$$  \hspace{1cm} (6)

The log function is called a barrier function and the $\mu$ in front of the log function is called the barrier parameter. Let’s say $x(\mu)$ is the minimizer of the function $f(x; \mu)$ then as $\mu$ approaches zero, $x(\mu)$ approaches the optimum $x$ for the original problem (viz: minimize $c^T x$). The set $\{x(\mu) : \mu > 0\}$ is called the central path. The main algorithm of the IPM finds a search direction that decreases the barrier parameter $\mu$ in subsequent iterations. The function $f(x; \mu)$ does not include the original constraint, which is $AX = B$. Then the problem can be summarized as “Find a minimum value of the objective function $f(x; \mu)$ which is subject to $AX = B$”. The problem can be solved using the standard Lagrange method.

$$\frac{\partial f(x; \mu)}{\partial x_j y_j} = \sum_{i=1}^{m} y_i \frac{\partial (A_{ij}x_j - b_i)}{\partial x_j y_j}$$  \hspace{1cm} (7)

The equation states that both vectors of the objective function and the constraints at an optimal point have the same direction - however, the sizes are different by constants $y_i$ times. The constants are called Lagrangian multipliers. In a standard LP problem which minimizes an objective function subject to equality constraints, the Lagrangian multipliers turn out to be the associated LP dual variables. The same concept is applied to the dual problem so the primal variables $x_i$ become Lagrangian multipliers in the dual problem Equation(8).

$$\frac{\partial f(y, z; \mu)}{\partial x_j y_j z_j} = \sum_{i=1}^{m} x_i \frac{\partial (A^T_{ij}y_j + z_i - c_i)}{\partial x_j y_j z_j}$$  \hspace{1cm} (8)

where $z_j$ are free variables added to the dual problem from the primal equality constraints. $c_i$ are the coefficients of the variables $x_i$ in the objective function. After computing Equation (7) and Equation (8), the following equations are obtained:

$$Ax = b, \quad (x > 0)$$  \hspace{1cm} (9)

$$A^T y + z = c, \quad (z > 0)$$  \hspace{1cm} (10)

$$xz = \mu e \quad (\forall e = 1)$$  \hspace{1cm} (11)

This system is called the KKT(Karush-Kuhn-Tucker) system with respect to $\mu$. In order for the Equations (7) and (8) to have an optimal minimum solution, there should exist vectors $x$, $y$, and $z$ satisfying the Equations (9)(10) and (11). Hence, the idea of the algorithm is that, given a $\mu$, solving both the primal and dual problem simultaneously makes the $\mu$ approach zero. That is the outline of the primal-dual algorithm for the IPM. The algorithm is implemented in GLPK [7] which we used for our modelling and analysis.
3.2. Tomogravity Method

The Tomogravity [4] method was developed by researchers at ATk&T. The method follows a two-step process. The first step involves generating an initial traffic matrix using the standard gravity model. The model assumes that the amount of traffic generated from a node or terminated at a node is proportional to the total size of flows from a node and inversely proportional to the squared distance between these two nodes. The inverse proportionality to squared distance of the gravity model is not considered in this study because “geographic locality is not a major factor in today’s Internet, as compared to ISP routing policies” [4]. Let the initial traffic matrix from the gravity model be \( M_g \). Suppose that the number of nodes in the network is \( n \), and \( C_j \) represents the column sum for column \( j \) in the traffic demand matrix. The value of \( R_i \) is the row sum for row \( i \) in this matrix. The row and column totals can be measured from the edge links using SNMP data. Then, the elements of this matrix can be obtained from the given row and column totals using the following equation:

\[
M_{(g)ij} = R_i \times \frac{C_j}{\sum_{k=1}^{n} C_k - C_i} \tag{12}
\]

The second step involves optimizing the initial traffic matrix obtained from the gravity model, using the least square method. The least square problem can be described as a quadratic problem as follows:

Minimise \( \| M - M_g \| \) \tag{13}

subject to the vector Equation (4). The least squares method searches for the best fitting set of values for the demands among many feasible solutions based on initial values and constraints. It calculates the error between the initial solution and other feasible solutions and then chooses one of the feasible solutions that produces the minimum error. In Equation (13), \( M \) is unknown, \( M_g \) (The initial traffic matrix from the gravity model, Equation (12)). In order to solve the vector Equation (13), the authors in [4] used MATLAB, whilst, for this paper, it was found that the commercial LP solver package CPLEX [8] could also be used.

4. EVALUATION METHOD

4.1. Simulation tools & Network Topology

A simulation program, using the \textit{C++} language, was built to simulate a network and to generate traffic distributions for the network. Two different network topologies (containing 8 nodes and 14 nodes) were created as shown in Fig. 3 and Fig. 4.

![Figure 3. An 8 node network.](image)

![Figure 4. A 14 node network.](image)

The fourteen node network had the same topology as that used in [2] and the eight node network was randomly created by the authors. CPLEX [8] was used for solving the quadratic
vector equation (13). CPLEX is a commercial product offering C and C++ libraries that solve linear programming and related problems. GLPK [7] (GNU Linear Programming Kit) was used to implement the IPM because the source code of GLPK is open to the public. This enables the algorithm for IPM to be analyzed from the source code.

To validate the IPM and the Tomogravity method, the RMSE (Root Mean Squared Error) and RMSRE (Root Mean Square Relative Error) were used, which provide an overall metric for the errors in the estimates. The RMSRE was calculated on the largest 75% of the flows as suggested in [4]. The reason is to protect the RMSRE from being dominated by small flows. The simulation was run 100 times in each case and the average RMSE and RMSRE were calculated.

4.2. A Synthetic Traffic Matrix

Ideally, real traffic matrices are required to compare the IPM to the Tomogravity approach. However, such information is generally not available to the public. Therefore, the authors used synthetic traffic matrices to perform the necessary comparisons. However, the problem of this approach is whether these traffic matrices closely represent real traffic matrices. In particular, comparing the LP method to the Tomogravity approach, this problem becomes more subtle because the Tomogravity method has only been verified in a real back-bone network. Therefore, a method was required to generate a synthetic traffic matrix which has similar properties to that of a back-bone network traffic matrix. Two different approaches were tried to generate such a traffic matrix.

Firstly, a uniform distribution was tried, because, traditionally the uniform distribution was the most common approach to create an element of the traffic matrix [9]. The traffic elements were generated from a random uniform distribution in the interval [100,500]. Fig. 5 shows the estimated traffic matrix from the Tomogravity method when the underlying distribution is a uniform. The solid diagonal line shows where the synthetic traffic matrix is estimated exactly and the dotted lines shows ±20% of the RMSRE of the estimated flows. The estimated points are spread roughly horizontally about the mean (300). Unfortunately, this result looked very different from the Tomogravity result in [4].

This is not so surprising as, after some consideration it is clear that a back-bone network is likely to have a “gravity distribution”. The interpretation of the gravity distribution is that the total traffic generated from a node is spread according to the basic gravity model. In a back-bone network, some (Points of Presence) POPs would attract or generate more traffic demand than other POPs because of attributes such as the size of the POP or the number of subscribers at
the POP locations. To simulate this “gravity distribution”, we modified the previous method [10]. In [10], node numbers were regarded as the size of the total flows generated from a node so that the volume of flow between two nodes would be decided by multiplying the node numbers of the two nodes. The problem with this approach is that the differences in node numbers are the same and the node numbers are in order. To solve this problem, the node numbers were decided by a uniform random number generated in the interval [1, total number of nodes] and then set as the weight \( \omega_i \) for each node \( i \).

\[
\omega_i = \chi_i \quad (i = 1, \ldots, N : \text{The number of nodes})
\]  

where \( \chi_i \sim \text{Uniform distribution [1, total number of nodes]} \) then the elements of the traffic matrix \( X_{ij} \) are generated as follows:

\[
X_{ij} = (\omega_i) \times (\omega_j) \times (R_{ij})
\]  

where \( R_{ij} \sim \text{Uniform distribution [(}\omega_i) \times (\omega_j), (2 \times (\omega_i) \times (\omega_j))] \), Equation(15) is used to generate the gravity distribution.

The weights, \( \omega_i \) and \( \omega_j \), actually represent the size of the total flows generated from a node (i or j) so that the volume of flow between these two nodes is decided by multiplying the weights of the two nodes. Therefore, a large flow exists between higher weighted nodes but these flows are scaled with a random factor based on a standard probability distribution. For this simulator, the node weights and the random factor have been drawn from the uniform distributions. Fig. 6 shows the result of the Tomogravity when a synthetic traffic matrix was generated by Equation 15. The RMSRE of the Fig. 6 is 14% which is quite similar to [4]. For the comparative study, we used Equation (15) to generate a synthetic traffic matrix.

5. RESULTS AND DISCUSSION

![Figure 7. Simplex method (RMSRE 130%)](image1)

![Figure 8. IPM (RMSRE 12%)](image2)

Figures 7 and 8 show that the two LP methods produce very different results in terms of estimating the distribution of the synthetic traffic matrix – although the optimal values of both methods are the same. The Simplex Method estimates many elements as zeros. In Fig. 7, many estimates lie on the X-axis. However the rest of the elements are over-estimated to compensate for the zero estimates. Only a few of the estimated points of the Simplex Method are between the \( \pm 20\% \) dotted lines while the estimates from the IPM are mostly between the \( \pm 20\% \) dotted lines. The RMSRE of the Simplex Method in Fig. 7 is 130%. This result explains why Medina et
al [2] reported that the errors with the LP method are so high, it could not be used in practical networks. However, clearly the LP approach with the IPM estimates the synthetic traffic matrix very accurately with 12% RMSRE in Fig. 8.

Figure 9. 8 node network IPM & TG

Figure 10. 14 node network IPM & TG

Table 1
The average RMSE and RMSRE

<table>
<thead>
<tr>
<th></th>
<th>8 nodes network</th>
<th>14 nodes network</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPM</td>
<td>1961.19%</td>
<td>2367.65</td>
</tr>
<tr>
<td>TG</td>
<td>2967.24%</td>
<td>3146.75</td>
</tr>
<tr>
<td>RMSRE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPM</td>
<td>12.0%</td>
<td>13%</td>
</tr>
<tr>
<td>TG</td>
<td>14.0%</td>
<td>14%</td>
</tr>
</tbody>
</table>

Table 1 shows the RMSE and the RMSRE of the IPM and the Tomogravity method for the two different topologies. The averages of the RMSREs were 12% and 13% for the IPM in the 8 and 14 node networks respectively, with 14% for the Tomogravity method in both networks. Although the IPM produced slightly less RMSRE and RMSE in each case in Table 1, it is hard to recognize the differences between Fig. 9 and Fig. 10. Therefore the method is most likely to give a better answer than any other models with the given distribution. However Fig. 9, Fig. 10, and Table 1 indirectly prove that the Tomogravity method does not have any special advantages over the IPM under the gravity distribution assumption.

Figure 11. 0% flows are zero

Figure 12. 20% flows are zero

Figure 13. 49% flows are zero
Table 2
The average of the RMSRE with some zero elements

<table>
<thead>
<tr>
<th>14 nodes network</th>
<th>0 % zero flows</th>
<th>12 % zero flows</th>
<th>67 % zero flows</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSRE</td>
<td>IPM</td>
<td>TG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12%</td>
<td>13%</td>
<td>12%</td>
</tr>
<tr>
<td></td>
<td>24%</td>
<td>25%</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>35%</td>
<td>35%</td>
<td>35%</td>
</tr>
</tbody>
</table>

Figures 11, 12, and 13 show what happens if some elements of a traffic matrix do not carry any traffic. This experiment was done because, the authors thought that the IPM might produce a larger error than the Tomogravity model in this case, as the number of zero elements increased in the real traffic matrix. However, both methods exhibit a similar behaviour. In Fig. 12 and Fig. 13, zero elements of the real traffic matrices are over-estimated; therefore, the other elements are under-estimated to compensate for the over-estimated values. Consequently, the solution points tend to go below the solid diagonal line as the percentage of zero elements increases. Also, the accuracy of both methods decreases as the percentage of zero elements increases according to Table 2.

Figure 14. Using hop counts for the weights of the objective function

Figure 15. Using the inverse product of the weight of an origin node and the weight of a destination node for the weights of the objective function

Table 3
The average of the RMSRE with the different objective function

<table>
<thead>
<tr>
<th>14 nodes network</th>
<th>Using hop count</th>
<th>Using node weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSRE</td>
<td>IPM</td>
<td>TG</td>
</tr>
<tr>
<td></td>
<td>12%</td>
<td>14%</td>
</tr>
<tr>
<td></td>
<td>26%</td>
<td>14%</td>
</tr>
</tbody>
</table>

Because the result of the LP approach is sensitive to the objective function, another set of weights was tried for the objective function in Equation (3). Equation (16) was derived from an idea that, in our LP problem, the less coefficient of an element in an objective function causes the more traffic demands to be allocated to the element.

\[ w_{ij} = \frac{1}{W_i \times W_j} \]  

(16)
where $w_{ij}$ is the weight of a flow (origin node $i$ to destination node $j$) in objective function, and $W_i$ and $W_j$ are the weight of a node $i$ and $j$ respectively. In Table 3, the accuracy of the IPM is changed from 12% to 26% when the different weights $w_{ij}$ of the objective function are used. Fig. 14 and Fig. 15 show the correlation between the weights of the objective function and the accuracy of the IPM. This result implies that the accuracy of the IPM can be improved choosing a proper objective function.

6. CONCLUSION

The LP method was regarded as a poor technique for traffic matrix estimation in [2]. However, we have shown from our study that the LP method based on the IPM algorithm could be a potential technique for estimating traffic matrices. Also, we have found that the Interior Point Method can produce similar or even slightly better results than the Tomogravity method.

The IPM does not require any pre-processing step such as obtaining an initial solution using the gravity model. However the method requires the choice of a proper objective function to obtain accurate results. The challenge ahead is to find an objective function that improves these results. It may depend on the network topology or the distribution of flows. In either case, it could lead to the possibility of “tuning” the objective function for different conditions.

As a minor contribution, we have introduced a new method to generate a synthetic traffic matrix which we believe has a similar distribution of traffic matrix elements as a back-bone traffic matrix. This method was extended from the previous method developed by the authors [10]. The method can be used to generate a traffic matrix for simulating a back-bone traffic matrix.

REFERENCES

10. S. Eum, J. Murphy, and R. Harris, TomoKruithof vs Tomogravity for Backbone Networks, ATNAC, Sydney, Australia, 2004.