

**Fundamental Properties of Synthetic O-D Generation  
Formulations and Solutions**

by

**Harinarayan Paramahamsan**

**Thesis submitted to the faculty of the  
Virginia Polytechnic Institute and State University  
In partial fulfillment of the requirements for the degree of**

**MASTER OF SCIENCE**

in

**Civil and Environmental Engineering**

**APPROVED:**

---

**Michel Van Aerde, Ph.D., Chairman**

---

**Hesham Rakha, Ph.D.**

---

**Antonio A. Trani, Ph.D.**

**January, 1999  
Blacksburg, Virginia**

# **Fundamental properties of Synthetic O-D Generation**

## **Formulations and Solutions**

**Harinarayan Paramahamsan**

**(ABSTRACT)**

Origin-Destination (O-D) matrices are required in order to model traffic routing behavior in networks. Direct techniques for collecting O-D information from home and roadside interviews have historically been utilized to estimate O-D tables. However, these techniques are not only very costly, labor intensive, and disruptive to trip makers, but traditionally also do not capture traffic peaking behavior, which is often required for traffic operational purposes. Consequently, more cost-effective indirect or synthetic O-D estimation techniques have been developed, and continue to be developed. They utilize readily available traffic volume counts to estimate the most likely O-D tables that may have generated the observed link counts.

This thesis describes the basic formulations that have been proposed to formulate and solve the static O-D problem synthetically using link flow observations based on Maximum Entropy techniques. As is the case with many mathematical solutions to engineering problems, a number of simplifying assumptions have been made in order to solve the synthetic O-D problem. Unfortunately, the descriptions of these simplifying assumptions are often not fully described in the literature, and in some cases, these assumptions are not mentioned at all. Furthermore, the literature fails to systematically demonstrate what impact these assumptions have on the final O-D table estimate. Therefore, this thesis utilizes simple hypothetical networks to;

1. Demonstrate and compare the two main types of synthetic O-D formulations, namely the trip and volume based formulations.
2. Track the O-D estimation procedure from its initial formulation to its final formulation, demonstrating all significant assumptions that have been made and the implications of these assumptions on the final solution.
3. Demonstrate to what extent the final O-D estimation formulation remains valid when these assumptions are invoked.
4. Test the applicability of some packages which implement the various formulations and solution techniques that are available.

## **Acknowledgments**

I would like to extend my sincere gratitude to Dr. Michel Van Aerde for serving as my advisor and helping me through the development of this thesis. Thanks also go to Dr. Hesham Rakha and Dr. Antonio Trani for not only serving as members of my committee, but for providing support and assistance whenever needed. Additional thanks to Dr. Michel Van Aerde for providing financial assistance through the Center for Transportation Research at Virginia Tech.

A very special thanks to my parents, Paramahamsan and Samyuktha, and my sister Nandini for their unflagging support during my academic years. None of the accomplishments I have achieved would have been possible without them.

Thanks to my friend Alexis Avgoustis, who I could always count on providing a distraction or lending an ear when I needed it most. Also thanks for the support and interest through it all.

Finally, thanks to all my fellow graduate students in the Transportation Infrastructure and Systems Engineering Division for providing encouragement and support towards the completion of this thesis and my graduate degree.

# Table of Contents

Acknowledgements .....	iii
Table of Contents .....	iv
List of Figures.....	vi
List of Tables .....	x
List of Symbols.....	xi
1 Introduction .....	1
1.1 Background.....	1
1.2 Need for O-D Trip Table.....	1
1.3 Thesis Objectives and Contributions .....	3
1.4 Thesis Organization .....	3
2 Literature Review.....	6
2.1 Introduction.....	6
2.2 Classification of Models.....	6
2.3 Solution Algorithms .....	11
2.4 Summary of Literature Review .....	12
3 Overview of Sub Problems .....	14
3.1 Introduction.....	14
3.2 Trip Distribution Vs. Synthetic O-D Generation .....	17
3.3 Multiple Solutions.....	19
3.4 Maximum Likelihood for O-D Matrices Based on Trips.....	21
3.5 Maximum Likelihood Based on Volume Counts .....	26
3.6 Presence of Flow Continuity .....	27
3.7 Other Approximations.....	31
3.8 Summary .....	33
4 Alternative Formulations of the Static O-D Problem .....	34

4.1	Introduction.....	34
4.2	Trip Formulations and Assumptions .....	34
4.3	Volume Formulations and Assumptions .....	41
4.4	Summary.....	43
5	Alternative Numerical Solution Techniques .....	44
5.1	Introduction.....	44
5.2	Double Step Approach .....	44
5.3	Single Step Approach.....	47
5.4	Solver Considerations and Attributes .....	49
	5.4.1 Computer Considerations .....	50
	5.4.2 Mathematical Considerations .....	51
5.5	Alternative Solvers and Packages.....	51
	5.5.1 Microsoft Excel Solver.....	51
	5.5.2 MATLAB .....	52
	5.5.3 QUEENS-OD .....	53
5.6	Summary.....	53
6	Comparison of Synthetic O-D Formulations.....	55
6.1	Introduction.....	55
6.2	Description of Test Networks.....	55
6.3	Formulations to be Tested .....	59
6.4	Effect of Scaling of Seed Matrix .....	61
6.5	Effect of Uniform Seed .....	72
6.6	Effect of Feasible Seed.....	87
6.7	Effect of Infeasible Seed .....	102
6.8	Effect of Flow Continuity.....	114
6.9	Summary.....	114
7	Conclusions and recommendations .....	116
7.1	Conclusions.....	116
7.2	Recommendations for Further Research .....	118
	References .....	119

Appendix A .....	123
Appendix B.....	126
Vita.....	133

# List of Figures

Figure 1.1: Flow Chart of Thesis Organization.....	4
Figure 3.1: 16-Link Network .....	17
Figure 3.2: 5-Link Network .....	21
Figure 3.3: Trips from A to C vs. Likelihood .....	24
Figure 3.4: 2-Link Network .....	32
Figure 3.5: Comparison of Optimal Solution; Including T and Excluding T .....	32
Figure 6.1: Five Link Network Configuration .....	56
Figure 6.2: Two Link Network Configuration .....	58
Figure 6.3: Optimal Solution vs. Scale Factor for the 5-Link Network Using Eq. 6.3 .....	64
Figure 6.4: Optimal Solution vs. Scale Factor for the 2-Link Network Using Eq. 6.3 .....	65
Figure 6.5: Optimal Solution vs. Scale Factor for the 5-Link Network Using Eq. 6.4 .....	66
Figure 6.6: Optimal Solution vs. Scale Factor for the 2-Link Network Using Eq. 6.4 .....	67
Figure 6.7: Optimal Solution vs. Scale Factor for the 5-Link Network Using Eq. 6.5 .....	68
Figure 6.8: Optimal Solution vs. Scale Factor for the 2-Link Network Using Eq. 6.5 .....	69
Figure 6.9: Optimal Solution vs. Scale Factor for the 5-Link Network Using Eq. 6.6 .....	70
Figure 6.10: Optimal Solution vs. Scale Factor for the 2-Link Network Using Eq. 6.6 .....	71
Figure 6.11: Optimal Solution for the 5-Link Network Using Equation 6.1 .....	74
Figure 6.12: Optimal Solution for the 5-Link Network Using Equation 6.2 .....	75
Figure 6.13: Optimal Solution for the 5-Link Network Using Equation 6.3 .....	76
Figure 6.14: Optimal Solution for the 5-Link Network Using Equation 6.4 .....	77
Figure 6.15: Optimal Solution for the 5-Link Network Using Equation 6.5 .....	78
Figure 6.16: Optimal Solution for the 5-Link Network Using Equation 6.6 .....	79
Figure 6.17: Optimal Solution for the 5-Link Network Using Equation 6.7 .....	80
Figure 6.18: Optimal Solution for the 5-Link Network Using Equation 6.8 .....	81
Figure 6.19: Optimal Solution for the 2-Link Network Using Equation 6.1 .....	82
Figure 6.20: Optimal Solution for the 2-Link Network Using Equation 6.2 .....	83
Figure 6.21: Optimal Solution for the 2-Link Network Using Equation 6.3 .....	84

Figure 6.22: Optimal Solution for the 2-Link Network Using Equation 6.7 .....	85
Figure 6.23: Optimal Solution for the 2-Link Network Using Equation 6.8 .....	86
Figure 6.24: Variation of Objective Function for Equation 6.1 for the 5-Link Network Under Feasible Seed Combination .....	88
Figure 6.25: Variation of Objective Function for Equation 6.2 for the 5-Link Network Under Feasible Seed Combination .....	89
Figure 6.26: Variation of Objective Function for Equation 6.3 for the 5-Link Network Under Feasible Seed Combination .....	90
Figure 6.27: Variation of Objective Function for Equation 6.4 for the 5-Link Network Under Feasible Seed Combination .....	91
Figure 6.28: Variation of Objective Function for Equation 6.5 for the 5-Link Network Under Feasible Seed Combination .....	92
Figure 6.29: Variation of Objective Function for Equation 6.6 for the 5-Link Network Under Feasible Seed Combination .....	93
Figure 6.30: Variation of Objective Function for Equation 6.7 for the 5-Link Network Under Feasible Seed Combination .....	94
Figure 6.31: Variation of Objective Function for Equation 6.8 for the 5-Link Network Under Feasible Seed Combination .....	95
Figure 6.32: Variation of Objective Function for Equation 6.1 for the 2-Link Network Under Feasible Seed Combination .....	96
Figure 6.33: Variation of Objective Function for Equation 6.2 for the 2-Link Network Under Feasible Seed Combination .....	97
Figure 6.34: Variation of Objective Function for Equation 6.3 for the 2-Link Network Under Feasible Seed Combination .....	98
Figure 6.35: Variation of Objective Function for Equation 6.4 for the 2-Link Network Under Feasible Seed Combination .....	99
Figure 6.36: Variation of Objective Function for Equation 6.7 for the 2-Link Network Under Feasible Seed Combination .....	100
Figure 6.37: Variation of Objective Function for Equation 6.8 for the 2-Link Network Under Feasible Seed Combination .....	101



Figure 6.38: Error vs. Frequency Plot for Equation 6.3 for the 5-Link Network..... 104  
Figure 6.39: Error vs. Frequency Plot for Equation 6.4 for the 5-Link Network..... 105  
Figure 6.40: Error vs. Frequency Plot for Equation 6.5 for the 5-Link Network..... 106  
Figure 6.41: Error vs. Frequency Plot for Equation 6.6 for the 5-Link Network..... 107  
Figure 6.42: Error vs. Frequency Plot for Equation 6.8 for the 5-Link Network..... 108  
Figure 6.43: Error vs. Frequency Plot for Equation 6.9 for the 5-Link Network..... 109  
Figure 6.44: Error vs. Frequency Plot for Equation 6.3 for the 2-Link Network..... 110  
Figure 6.45: Error vs. Frequency Plot for Equation 6.7 for the 2-Link Network..... 111  
Figure 6.46: Error vs. Frequency Plot for Equation 6.8 for the 2-Link Network..... 112  
Figure 6.47: Error vs. Frequency Plot for Equation 6.9 for the 2-Link Network..... 113  
Figure A.1: X vs. Ln(X)..... 123

## List of Tables

Table 3.1: Complementary Nature of Synthetic O-D Process and Trip Distribution.....	18
Table 3.2: Comparison of Trip Distribution vs. Synthetic O-D Generation .....	19
Table 3.3: Observed Link Flows.....	22
Table 3.4: General Matrix.....	22
Table 3.5: Potential Matrix .....	22
Table 3.6: Possible Solutions for the 5-Link Network .....	25
Table 3.7: Matrix of Fraction of Traffic Obtained from Previous Survey .....	25
Table 3.8: Seed Matrix Based on Fraction of Traffic .....	25
Table 3.9: Link Volume on the 2-Link Network.....	33
Table 6.1: Link Volume on the 5-Link Network.....	56
Table 6.2: Feasible O-D Trip Matrix for the 5-Link Network.....	57
Table 6.3: Solution Set for the 5-Link Network.....	57
Table 6.4: Link Volume on the 2-Link Network.....	58
Table 6.5: Feasible O-D Trip Matrix for the 2-Link Network.....	58
Table 6.6: Solution Set for the 2-Link Network.....	59
Table 6.7: Comparison of Optimal Solution for Uniform Seed.....	73
Table 6.8: Summary of Results.....	115

## List of Symbols

Symbol	Description
$T_{ij}$	The actual but often unknown number of trips going from origin $i$ to destination $j$ .
$t_{ij}$	Initial estimate of number of trips going from origin $i$ to destination $j$ sometimes referred to as the seed matrix.
$T = \sum_{ij} T_{ij}$	Total number of trips traversing the network (The summation of all trips across all origin/destination combinations).
$t = \sum_{ij} t_{ij}$	Total number of trips traversing the network for the seed matrix (The summation of all trips across all origin/destination combinations within the seed matrix).
$T_k$	Number of trips available in the $k^{\text{th}}$ cell of a $p \times p$ O-D matrix; $k = i + (j - 1) \times p$ .
$T_n = T - \sum_{k=1}^{k=n-1} T_k$	Total number of trips available to be considered in selecting $T_{ij}$ trips.
$P_{ij}^a$	The fraction of trips going from origin $i$ to destination $j$ that use link $a$ .
$v_a = \sum_{ij} t_{ij} P_{ij}^a$	Volume on link $a$ associated with seed matrix $t$ .
$V_a = \sum_{ij} T_{ij} P_{ij}^a$	Estimated volume on link $a$ .
$V'_a$	Complementary traffic flows to $V_a$ which satisfy flow continuity.