Analysis of the Distribution of Traffic Density Using the Ideal Flow Method and the Principle of Maximum Entropy

Analysis of the Distribution of Traffic Density Using the Ideal Flow Method and the Principle of Maximum Entropy

R.W. Gardon
Asia Pacific College
3 Humabon Place, Magallanes
Makati City 1232
roselleg@apc.edu.ph

K. Teknomo
Ateneo de Manila University
Katipunan Avenue, Loyola Heights
Quezon City 1108
kteknomo@ateneo.edu

ABSTRACT

The persistence of the problem of traffic congestion and the increase in the demand for mobility has led to the widespread use of navigation apps. People not only have to get to their destination, they also want to get there in the soonest time possible. Thus, in this paper, a network traffic model called the ideal flow model is used to analyse a way of distributing traffic density across all road networks in the best way possible. Since the distribution of traffic density is one of the bases of traffic prediction, one which involves uncertainty, the principle of maximum entropy is utilized to determine the properties / behaviour of travel time based on the distribution.

CCS CONCEPTS

•Computing Methodologies → Modeling and simulation;

KEYWORDS

Ideal Flow Model, Maximum Entropy, Traffic Density Distribution

1 INTRODUCTION

Transportation is a derived demand [7]. People move because they need to get to some destination. With the increase in the demand for mobility, people do not only need to get to where they want to go, they also want to be at their destination in the fastest time possible. This led to the need for mobile navigation apps such as Google Maps and Waze, among others, that provide real-time traffic information. These apps help users choose which is the best possible route; and people usually base their decisions on the route that can get them to their destination in the shortest travel time. The information that users get from these apps include, but not limited to, congestion, road accidents, and road closures. However, the traffic information obtained from these apps only play a passive role; that is, the knowledge of the current traffic information does not decrease traffic density [7], and thus, does not actually alleviate traffic congestion. One of the goals of obtaining traffic information is to get users to their destination in the shortest amount of time possible. This can be done by applying a traffic model that distributes the amount of traffic density across all road networks in the best possible way. Traffic flow prediction is one of the bases on how traffic density will be distributed, and thus, entails uncertainty. In this paper, a network traffic model called the ideal flow model [8] is used to analyze a way of distributing traffic density based on traffic prediction. Moreover, since uncertainty is involved, the principle of maximum entropy from information theory [4] is utilized to determine the properties / behavior of travel time based on the distribution. The rest of the paper is structured as follows. In the next section, related works on traffic distribution are discussed. This includes an overview of the ideal flow model and a description of the principle of maximum entropy. After that, computations involved in the analysis of traffic density distribution using the ideal flow model and the principle of maximum entropy is presented. Finally, conclusions are drawn in the last section.

2 RELATED WORKS

2.1 Ideal Flow Model

Teknomo [8, 9] defined the ideal flow as the "aggregation of the trajectories of results in random walk in a uniformly distributed flow over space and time in a network graph". In other words, the ideal flow model is based on the random walk of multi-agents in a directed network graph [8, 9]. This model aims to determine the ideal distribution of traffic flow in a network, and this phenomenon happens when the "flow is distributed uniformly over space and time" [8]. This model can be applied to generally any probability distribution using Markov chain. In a Markov chain, in a very long random walk, the probability that an agent ends at some vertex in a network graph is independent of the agent's starting point [3]. Thus, in an ideal flow model, the network graph must necessarily be strongly connected [9]. The framework for traffic assignment used in the development of the ideal flow model is illustrated in Figure 1. The details of the framework and an application example

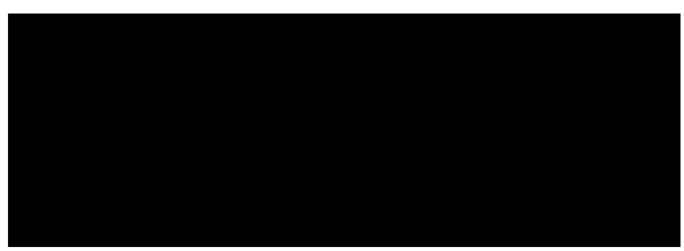

Figure 1: Ideal Flow Model Framework [10]

are in [10], but a reference for the notations will be included here since these are required in the next section.

Suppose a network structure is given with adjacency matrix $\mathbf{A} = [a_{ij}]$. A link distance matrix $\mathbf{L} = [l_{ij}]$, a link capacity matrix $\mathbf{C} = [c_{ij}]$, and a link maximum speed matrix $\mathbf{U} = [u_{ij}]$ can be constructed using the same matrix structure as **A**. The stochastic matrix $\mathbf{S} = [s_{ij}]$ is created such that

$$s_{ij} = \frac{c_{ij}}{\sum_{j=1}^{n} c_{ij}} \tag{1}$$

where n is the number of rows of the matrices. The ideal flow matrix is defined as $\mathbf{F} = \left[f_{ij}\right] = \left[\mathbf{f_1^T} \quad \mathbf{f_2^T} \quad \dots \quad \mathbf{f_n^T}\right]^T$, where

$$\mathbf{f_i^T} = \boldsymbol{\pi}_\mathbf{i}^T \mathbf{s_i^T} \qquad \text{for } i = 1, 2, \dots, n. \tag{2}$$

The matrix $\boldsymbol{\pi} = \left[\boldsymbol{\pi}_i\right]$ is the node probability distribution derived when the following optimization problem is solved.

$$(\mathbf{S}^T - \mathbf{I})\boldsymbol{\pi} = 0$$

subject to the constraints

$$\begin{aligned} \mathbf{j}^T \boldsymbol{\pi} &= 1 \\ \boldsymbol{\pi} &\geq 0 \end{aligned}$$

where $\mathbf{j}^T = \left[1 \quad 1 \quad \dots \quad 1\right]$.
When the ideal flow matrix is obtained, the flow/capacity ratio matrix

$$\mathbf{W} = \left[w_{ij}\right] = \begin{cases} \frac{f_{ij}}{c_{ij}}, & c_{ij} \neq 0 \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

can be derived.
To determine the travel time, Greenshields' [2] cost function is applied. The travel time formula is, thus, as followed:

$$\mathbf{T} = \left[t_{ij}\right] = \frac{2l_{ij}}{u_{ij}(1 + \sqrt{1 - w_{ij}})} \tag{4}$$

2.2 Principle of Maximum Entropy

The idea of information entropy [6] was introduced by Shannon in 1948 while Jaynes [4] introduced the principle of maximum entropy as a means of statistical inference. Entropy refers to a measure of uncertainty. The following definition is from [1].

Definition 2.1. Given a discrete probability distribution p on the countable set $\{x_1, x_2, \dots\}$ with $p_i = p(x_i)$, the *entropy of p* is defined as

$$H = -\sum_{j \geq 0} p_j \log_2 p_j \tag{5}$$

The principle of maximum entropy, which is credited to Jaynes [4], as stated in [5], is the following:

> "Out of all probability distributions consistent with a given set of constraints, the distribution with maximum uncertainty should be chosen."

The distribution of traffic density, therefore, that must be chosen, based on the above statement, is the one with the highest entropy. A distribution with the highest entropy is supposedly the one that gives the least of surprises in the predictions that it creates [1].

3 COMPUTATION AND ANALYSIS

3.1 Numerical Examples

3.1.1 Entropy of a Graph. Consider the strongly connected graphs in Figure 2 and Figure 4 with the indicated capacities.

Figure 2: Generalized Ideal Flow Example, with Capacities

Example 3.1 (Generalized Ideal Flow). Refer to figure 2. In a generalized ideal flow, the inflow to the node is equal to the outflow from the node.
To compute for the network entropy, refer to figure 3 for the probability of each link.

Figure 3: Generalized Ideal Flow Example, with Probabilities

The entropy of figure 2, therefore, is

$$H = -(\frac{4}{9}\log_2\frac{4}{9} + \frac{5}{9}\log_2\frac{5}{9} + \frac{3}{5}\log_2\frac{3}{5} + \frac{2}{5}\log_2\frac{2}{5} + 2(1\log_2 1))$$
$$H \approx 1.96$$

Example 3.2. Refer to figure 4. In a standardized ideal flow, the capacity of each outflow link is equal to the inflow to a specific node divided by the outdegree of that node.

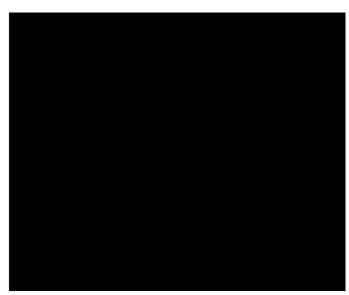

Figure 4: Standardized Ideal Flow Example, with Capacities

The probability of each link is indicated in figure 5.

Figure 5: Standardized Ideal Flow Example, with Probabilities

The entropy of figure 4 is

$$H = -(4(\frac{1}{2}\log_2\frac{1}{2}) + 2(1\log_2 1))$$
$$H = 2$$

3.1.2 Ideal Flow Model for Traffic Assignment. Suppose there is a road network as in figure 2, and that each edge weight times 1000 corresponds to the respective capacity. Furthermore, in this example, assume the following:

(1) The base maximum speed for each edge is 30 kilometers per hour for the first lane, and that there is an additional speed of 10 kilometers per hour for each additional lane.
(2) Each edge has length 500 meters.

The preceding network specification can be represented as matrices using the definitions presented in section 2.
The capacity matrix corresponding to figure 2 is

$$C(vehicle/hour) = \begin{array}{c} a \\ b \\ c \\ d \end{array} \begin{bmatrix} 0 & 5000 & 0 & 4000 \\ 0 & 0 & 2000 & 3000 \\ 9000 & 0 & 0 & 0 \\ 0 & 0 & 7000 & 0 \end{bmatrix}$$

Based on the preceding assumptions, the link maximum speed matrix and the link distance matrix are

$$U(km/hour) = \begin{array}{c} a \\ b \\ c \\ d \end{array} \begin{bmatrix} 0 & 70 & 0 & 60 \\ 0 & 0 & 40 & 50 \\ 110 & 0 & 0 & 0 \\ 0 & 0 & 90 & 0 \end{bmatrix}$$

and

$$L(km) = \begin{array}{c} a \\ b \\ c \\ d \end{array} \begin{bmatrix} 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 \end{bmatrix}$$

respectively.
Deriving the stochastic matrix S from C gives

$$S = \begin{array}{c} a \\ b \\ c \\ d \end{array} \begin{bmatrix} 0 & \frac{5}{9} & 0 & \frac{4}{9} \\ 0 & 0 & \frac{2}{5} & \frac{3}{5} \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Now, solving for the node probability distribution $\boldsymbol{\pi}$, we obtain

$$\boldsymbol{\pi} = \begin{bmatrix} \frac{3}{10} \\ \frac{1}{6} \\ \frac{3}{10} \\ \frac{7}{30} \end{bmatrix}$$

Then, solving for the ideal flow matrix F results to

$$F = \begin{array}{c} a \\ b \\ c \\ d \end{array} \begin{bmatrix} 0 & \frac{5000}{3} & 0 & \frac{4000}{3} \\ 0 & 0 & \frac{2000}{3} & 1000 \\ 3000 & 0 & 0 & 0 \\ 0 & 0 & \frac{7000}{3} & 0 \end{bmatrix}$$

Some interesting notes, due to Teknomo [8], about the matrix F:

- The matrix is pre-magic; that is, the vector row sum is equal to the transpose of the vector column sum.
- The vector row sums are equal to the elements of $\boldsymbol{\pi}$.

Since F and C are already available, the flow/capacity ratio can now be computed, and it is as follows:

$$W = \begin{array}{c} a \\ b \\ c \\ d \end{array} \begin{bmatrix} 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 \end{bmatrix}$$

Finally, by applying equation (4) to get the travel time matrix T,

$$T(hour) = \begin{array}{c} a \\ b \\ c \\ d \end{array} \begin{bmatrix} 0 & \frac{478}{1013} & 0 & \frac{485}{881} \\ 0 & 0 & \frac{782}{987} & \frac{691}{1046} \\ \frac{324}{1079} & 0 & 0 & 0 \\ 0 & 0 & \frac{396}{1079} & 0 \end{bmatrix}$$

3.2 Analysis

Referring to the last statement in section 2, the distribution that must be used in a traffic assignment model is the one with the highest entropy. One such distribution that is actually utilized in the ideal flow model is the uniform distribution. It is proven in [1] that the distribution that gives the maximum entropy is the uniform distribution. To determine the relationship of maximum entropy with travel time, different cases are considered.

Case 1. **L**, **U** are constant (i.e. nonzero elements inside **L** are the same, nonzero elements inside **U** are the same)

Case 1.1. If **W** is constant, travel time **T** becomes constant. That is, for all links/edges, the value of the travel time is equal to the average travel time. Moreover, it can easily be shown through simple calculus that the maximum value of T is obtained when **W** = 1. This can also be deduced from Fig 6.

Figure 6: Graph of T

Case 1.2. If **W** is not constant, then the values in T will have the same behavior as **W**; that is, if the flow/capacity ratio increases, the travel time increases as well.

Case 2. Either **L** or **U** is constant, but not both
Whether **W** is a constant or not, the travel time will not be fixed anymore.

Case 3. **L**, **U**, **F**, **C** constant
The result here is the same as in Case 1.1.

Case 4. **L** and **U** constant; either **F** or **C** is constant, but not both
The result here is the same as Case 1.2.