Smart Driving: Reducing Pollution by "Greenest" Paths

by Benedetto Piccoli (Math Dept and CCIB, Rutgers University – Camden)
Gene Fiorini (DIMACS, Rutgers University)

Module summary

The impact of pollution produced by vehicular traffic is one of the most important factors affecting public health. According to an article published by R. J. Laumbach and H. M. Kipen in the *Annals of the New York Academy of Sciences*, epidemiological studies "have linked exposure to traffic-related air pollutants to increased respiratory and cardiovascular morbidity and mortality." The phenomenon is of particular importance in metropolitan areas worldwide.

Transportation is a major consumer of energy. According to the 2009 Annual Energy Review of the U.S. Energy Information Administration transportation accounts for 27% of the share of energy used in the United States. Advances in new technologies for info-mobility provide consistent sources of data on traffic load for most urban areas. It is possible to take advantage of these advances in info-mobility in order to improve the efficiency of energy usage by transportation.

This module focuses on choosing the "least polluting" path when driving from a given starting position A to a given arrival position B. Students may use data available on the web to estimate the pollution effect.

Informal description

The module introduces some basics of vehicular traffic modeling such as the concepts of velocity vs density. A simplified model for fuel consumption is then developed as function of velocity and acceleration. This model will be used together with Google maps, MapQuest or other online sources to analyze data to decide which path will be the "greenest."

Target Audience

The target audience is students with some basic knowledge of optimization techniques. This includes students in one of the Calculus courses, a discrete mathematics course, or even computer science. For the advanced optimization techniques in the appendix, operational research or a graph theory course.

Prerequisites

Precalculus. For the advanced appendix section optimal paths over graphs.

Mathematical Fields

Graph theory, Optimization.

Appplication areas

The results of the module could be applied to produce more efficient travel by computing "greenest" paths in a heavily traveled area. If applied on large scales this would substantially contribute to reduction of pollution due to transportation.

Goal and objectives

The final goal is to enable road network users to diminish the pollution due to their use of the network.

The main objective is to provide students with math and modeling skills enabling them to choose the "greenest" path, but even to project the programming of applets for computing "greenest" paths.

Tehcnology/Software needs.

Web browser to access Google maps, Map Quest or other sources. Matlab for the more advanced section.

Instructor Notes: To make the module go more smoothly, it is suggested that the instructor present the module to the students in advance of covering it in class with instructions to read it over in preparation for completing the module in the next class period.

The module can be partitioned into several sections, some of which can be treated as "optional" if time constraints are a concern and to maximize the benefits to students:

1. Preliminary Class Discussion: Variations to defining a "greenest" path – This section can be treated as optional and should only be done if the instructor has sufficient time and thinks it will benefit the students to have a more extended discussion on what is meant by "greenest" path. If class time is used, it should be limited to a few minutes to allow maximum time to review the main model.

2. Glossary of Terms: This section should be assigned as advanced reading. Most students should be familiar with these terms or should be able to quickly pick them up on their own. Class time would be better spent on discussing the main model.

3. The Model: Two versions of the main model are given – solution (a) and solution (b). The majority of class time should be concentrated on solution (a). Solution (a) is a good demonstration of the linear model and will give the students a good understanding of "greenest" path. Solution (b) is a more complex version of solution (a) and could be treated as "optional," especially if time is a consideration. Fully developing solution (a) should be approximately 30 to 40 minutes.

4. Extensions at the end are definitely optional sections. Students who demonstrate and ability and an interest could be encouraged to work on the extensions on their own or for extra credit.

The Module

<u>Introduction</u>

There are a number of web mapping service applications, such as Google Maps and MapQuest that now provide a number of functions that give additional information about travel routes. If someone is planning a trip from a starting position A to a given destination position B, the traveler now has access to useful travel information such as distance traveled, alternate routes, and expected travel time that could improve the traveler's trip in a number of ways. When planning to travel from a given starting position A to a given destination position B, is it possible to use the information provided by a web mapping service application to solve the problem of finding the "greenest" or least polluting path from A to B? The term "green" is often used today to refer to a method or a process that is environmentally friendly or concerned with behaving in a sustainably-conscience way. The phrase "greenest" path within the context of this activity could have several interpretations. Any number of factors associated with traveling could have an impact on how "green" a trip is measured to be.

This module will begin by exploring some factors that affect the "green-ness" of a trip before identifying a definition for greenest path that will be used for this activity. Once a common definition for greenest path is established, the module introduces some basic terminology and properties of graph theory that are used to develop a model for measuring the greenest path by calculating a "green" factor along the route traveled. This model is then extended using the same basic graph theoretic principles to develop variations of the model which are then compared to the original model. The models presented are linear model. This "simplified" approach illustrates what could be accomplished in a relatively short period of time using accessible methods and tools. The module concludes with some optional, additional material that gives students an opportunity to apply the methodology using more advanced mathematical techniques.

Instructor Notes: Instructors and students are encouraged to explore other interpretations of "greenest path" and compare if this has any impact on which route the model identifies as greenest. Once students are familiar with the basic model, they should be encouraged to develop their own model based on reasonable assumptions of their own to determine if a different model identifies a different greenest path.

Preliminary Class Discussion: There are many factors affecting which path could be identified as the "greenest" or optimal. Before reading ahead, spend some class time discussing possible factors. Are there some factors that would have a greater impact than others on determining the optimal path?

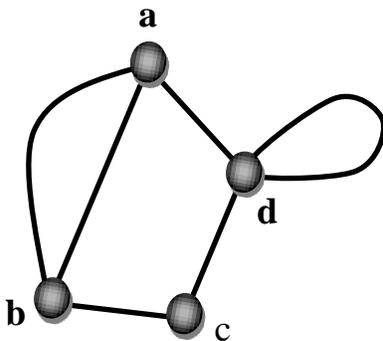For this activity the factors that will be considered are:

- Length of paths
- Expected travel time for each path
- Acceleration contribution over the path

Data for these three factors are readily available (or can be computed) from Google maps or MapQuest for a given set of paths. Using optimization over graphs it is possible to consider a larger set of paths.

Glossary of Terms

There are a few terms that will be useful when working through the activities in this module. Readers who are familiar with these terms can proceed directly to the next section that introduces the model.

A **graph (or network) G** is a set of at least one **vertex**, $V(G) = \{v_1, v_2, \ldots, v_n\}$, commonly represented as points, and a set of **edges** $E(G) = \{e_1, e_2, \ldots, e_n\}$, commonly represented as a line or curve having a vertex at each end. The edge connecting vertices **u and v** is written **uv**.

There are many different types of graphs. The most common type of graph, and the one used in this activity, is one in which at most one edge joins two distinct vertices. Because the vertex-edge connections do not depend on position, a graph has many visually different representations.

A **path** in a graph is a sequence of distinct vertices connected by edges. The **length** of a path is the number of edges (or one less than the number of vertices) contained in the path.

A **mathematical model** is an abstract representation that uses mathematical language and concepts to describe the behavior of a system. Mathematical models are used in the natural sciences (physics, biology, earth science), engineering sciences (computer science, artificial intelligence), and the social sciences (economics, psychology, sociology, political science). A model can be used to help explain different components of a system and make predictions about its future. Mathematical models can take many forms: dynamical systems, statistical systems, differential equations, graphical systems, and game theoretic models, as well as many others.

An **optimization problem** is the challenge of finding the best solution out of all feasible solutions to that problem. There are two categories of optimization problems: discrete and continuous. A discrete optimization problem is sometimes called a **combinatorial optimization problem**.

A **road network** is a set of roads connected by junctions. Mathematically this can be represented by a *graph* with edges representing roads and vertices representing junctions. A car moving over the network travels over links (*edges* of the graph) with a given **velocity** assigned to each link. The phrase **variation of velocity** (or **velocity**

variation) in this module refers to the changing velocity value from link to link (edge to edge) along a path.
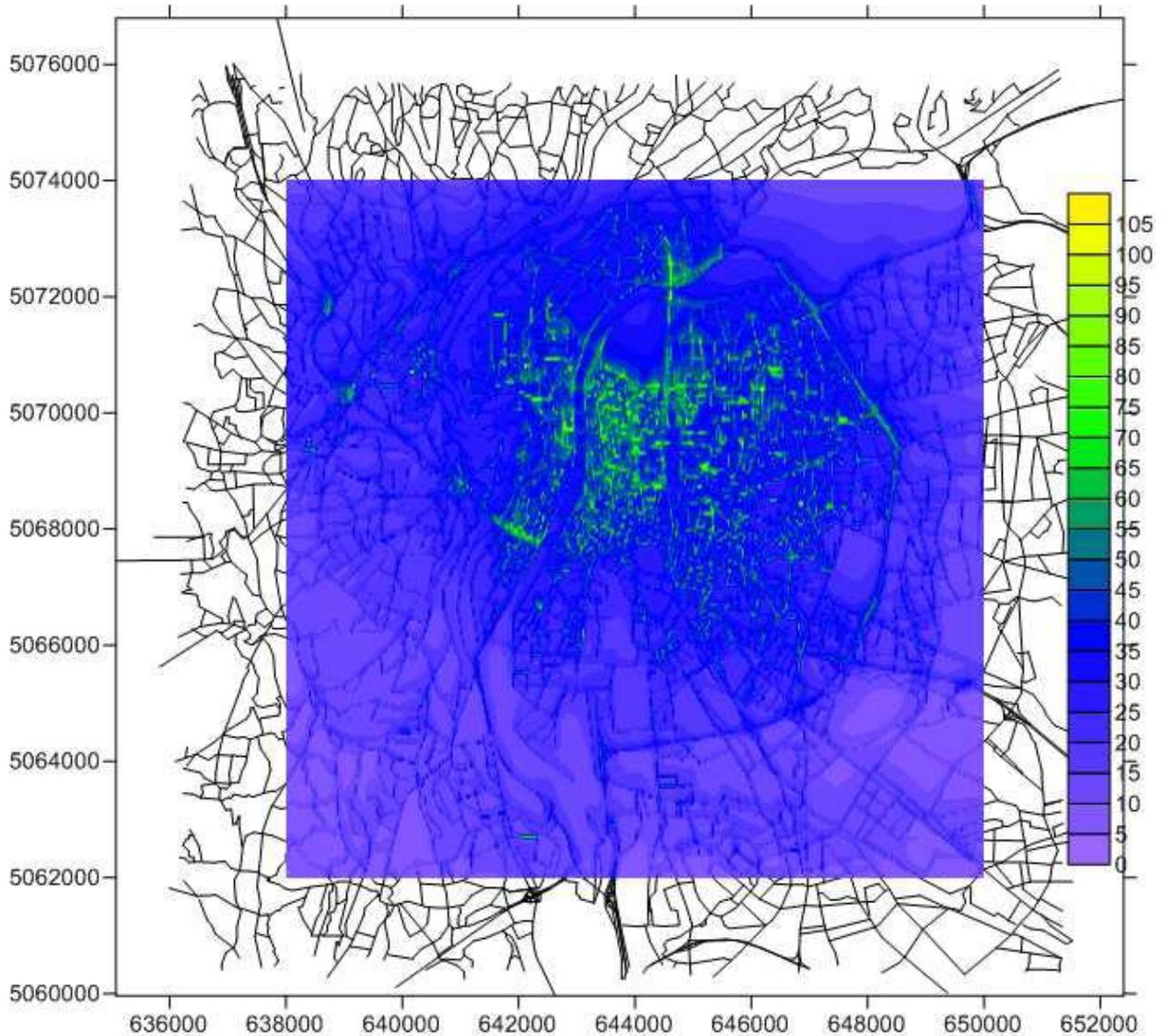
The **fuel consumption** is the number of gallons used to complete the travel and the **pollution** produced by the travel is the total quantity of particles emissions during the travel.

The Model

The image in Diagram 1 is a simulation of a city network (roadways) and the relative pollution due to traffic. The color scale measures number of fine particulate matter particles due to the pollution effect directly related to the traffic load over the network.

Diagram 1

Question: Using the scale to the right of the diagram:

1) Which color suggests the heaviest amount of traffic? Which color suggests the least amount of traffic?
2) Based on this information, where in the city would a single car be expected to have the greatest fuel consumption? The least?

Using the relative pollution due to traffic, it is possible to determine the fuel consumption of a single car given where the car is located in traffic. However, there are better ways of determining the fuel consumption of a single car and thus the relative polluting effect. This can be done by taking into consideration other factors such as length of path, expected travel time along the path, and acceleration.

Problem: Given a starting position A over the network and destination position B, determine the "greenest" path, i.e. the least polluting one, by comparing the fuel consumption along different paths from A to B, given the length of the path from A to B, the expected travel time along that path, and the acceleration along that path.

To solve this problem it is formulated as a mathematical optimization problem in the following way:

> A road network can be represented as a graph $= (V, E)$, where the set of vertices $V = \{v_1, v_2, \ldots, v_n\}$ represent locations along the network where the commuter has the option to change their driving "conditions" (such as direction at intersections, traffic circles, interstate interchanges; speed when traffic volume changes; etc.) and two vertices are joined by an edge if there is a direct route between the two vertices without the option of changing paths along that portion of the route. .Let A be a starting vertex and B a destination vertex. Assuming that a pollution cost can be assigned to each path, find a path connecting A to B with minimal cost.

There are several approaches to solving this optimization problem:

Solution a).
If the "green" cost of each complete path is known, then all paths from A to B can be considered and the minimal path can be determined by comparison.

Solution b).
Sometimes it is possible to assign "green" costs to each edge along a path. Then the cost of each path is the sum of the costs of edges. To find an optimal path with minimal green costs, it is possible to use well know algorithms from Operations Research or Graph Theory. These algorithms could be evaluated with the use of a Matlab package.

First consider **solution a)**. There exist data and models to compute the cost of each path from A to B.

Recent research provides very reliable models for measuring fuel consumption (FC) for a given route of travel (video of a presentation at MIT available, see References). A simplified formula is given by:

$$FC = IDFC + a \cdot D + b \cdot TT + c \cdot V, \tag{1}$$

where

> $IDFC$ is the fuel consumption due to "idle time", for example waiting time at traffic light;

D is the distance covered;

TT the travel time;

V the variation in velocity; and,

a, b, c are appropriate coefficients.

Data on the quantity $IDFC$ are difficult to find. Therefore, for the purposes of this activity, the value of $IDFC$ will assumed to be constant over every A-B path. As a result, its coefficient can be taken to be 1.[1]

Computing the other quantities, $D,\ TT,\ V$, in formula (1) begins with estimating their respective coefficients $a, b, and\ c$. The coefficient values have been estimated through experimental observation and represent the contribution of each corresponding quantity relative to the corresponding standard unit of measurement. First, note that by repeated experimental measurements it has been determined that the contribution of each quantity ($D,\ TT,\ V$) is approximately 5%, 40% and 35% respectively.

These percentages need to be represented relative to each corresponding unit of measurement. For simplicity each corresponding relative unit of measurement is typically taken to be one of the following values: 1mile (for distance), 1 minute (for travel time), 10 miles/h (for velocity). Moreover, on average it will be assumed that typical average fuel consumption is 1 gallon for every 20 miles driven. Therefore, with these standard assumptions, the coefficients can be calculated to be[2]:

$$a = 0.05 \cdot \left(\frac{1}{20}\right) \frac{gallons}{mile} = 0.0025\ gal./mile$$
$$b = 0.4 \cdot \left(\frac{1}{20}\right) \frac{gallons}{minute} = 0.02\ \frac{gal}{min} \qquad (2)$$
$$c = 0.35 \cdot \left(\frac{1}{10}\right) \frac{gallons}{\left(\frac{miles}{hr}\right)} = 0.0175\ \frac{gallons}{\left(\frac{miles}{hr}\right)}$$

For each given path it is now possible to compute FC.

The quantities $D,\ TT,\ and\ V$ can be computed using information available on most web mapping systems. For instance, FC can be calculated using Google maps or Map Quest for most areas. For specific areas, it may be possible to use additional advanced monitoring for traffic. For the San Francisco Bay data from the Millennium project

---

[1] If $IDFC$ is not assumed to be constant, alternatively it can be calculated by counting the number of traffic lights and stops along the path in question. This will be explored in the optional section at the end of the activity.
[2] Depending of different time zones under consideration or specific vehicles, the coefficients may be corrected or changed.

(traffic.berkeley.edu) could be used. For Italy and some European cities the data from Octotelematics (http://traffico.octotelematics.it/index_en.html) could be used.

Using Google maps for data

Class Activity: What would George Washington have done if he had access to a web mapping system?: Driving from Independence Hall, Philadelphia, PA to the Valley Forge Visitor Center, Valley Forge, PA.

1) Go to Google maps[3] and type in the address for the Valley Forge Visitor Center (destination location B) into the Google search bar:

> 1400 North Outer Line Drive
> King of Prussia, PA 19406

Click on the map that appears in the search.

2) Click on the directions link (note that your address will appear in the 'B' address bar to the left of the map. In the 'A' address bar type in "Independence Hall, Walnut Street, Philadelphia, PA

3) Press 'Enter." A map with route outlined on it will appear to the right. On the left Google maps usually proposes two or three route alternatives. With each alternative Google maps provides

- Travelling distance which is D;
- Expected travel time which is TT.

Record the distance and expected travel time for a route given.

Note: In case Google only provides one option, additional alternatives can be found by clicking and holding on the given route that appears in the map and dragging it to another road. The directions, distance and expected travel time on the left will adjust to match the new route.

4) The only quantity left to calculate in equation (1) is the velocity variation, V. In the upper right hand corner of the map click on the "traffic" option. The "traffic" option provides a pictorial description of the traffic data. Up to four colors appear on the road map: black, red, yellow, and green. The colors represent traffic volume: black representing the heaviest volume of traffic through green representing the lightest volume of traffic. This information can be used to estimate the variation in velocity, V. Using some reasonable assumptions, a representative velocity can be estimated for each color. The first assumption will be that the velocities association with each color will be equidistant; the least velocity associated with black, and the greatest velocity associated with green. Furthermore, 40 mph will

---

[3] This activity could also be done with MapQuest or a similar web mapping system.

be used as an upper bound for the maximum velocity. These assumptions are consistent with previous estimates given for equation (1). Based on these assumptions, segments of black will be assigned a velocity of 5 m/h, red segments will be assigned a velocity of 15 m/h, yellow segments will be assigned a velocity of 25 m/h, and green segments will be assigned a velocity of 35 m/h. From this information we can compute the variation of velocity for the path in the following way:

$$\sum_{i=1}^{n-1} |Color(i+1) - Color(i)|, \tag{3}$$

where $Color(i)$ is the numerical value of velocity corresponding to the ith segment link and the sum runs over all the segment links composing the A-B path.

Example: Calculate the variation of velocity from the intersection of Via Ancona and Via Nomentana to the intersection of Via San Martino della Battaglia and Piazza Indipendenza in Rome, Italy (identified on the map below).

To calculate the variation in velocity, note that there are four links in the path between the two intersections: 1) Via Nomentana (red – 15mph) to 2) Corso

D'Italia (green – 35mph) to 3) Viale Castro Pretorio (yellow – 25mph) to 4) Via San Martino della Battaglia (red – 15mph). Therefore,

$$v = |35 - 15| + |25 – 35| + |15 – 25| = 20 + 10 + 10 = 40$$

Calculate the variation of velocity for the path identified by Google maps (or created by moving the path to another route). Note that answers will vary depending on what time of day, what day of the week, and which routes are chosen.

Instructor Notes: Instructors may want to alter the assumptions used to calculate the variation of velocity for each color segment. For example, if only interstate routes are being considered, assuming a maximum velocity of 40 mph does not make sense. If time permits, it might be interesting to have students discuss what other assumptions would make sense and under what conditions. Students could be asked to predict what would happen if the assumptions were altered, and then compare the results with their predictions. Students could also be asked if changing assumptions about the variation of velocity would change the answer of which path is "greenest."

5) All the data for computing the pollution cost (fuel consumption) for the identified A-B path are now available. Plug in all the calculated values ($D, \ TT, \ V, a, b, c$) into expression (1) to calculate the pollution cost for the A-B path.

   Question: Note that the value of $IDFC$ was not calculated. Explain why we do not need to calculate the value of $IDFC$ in order to determine which path is optimal. *Answer:* $IDFC$ is constant. As such it does not affect the maximum value.
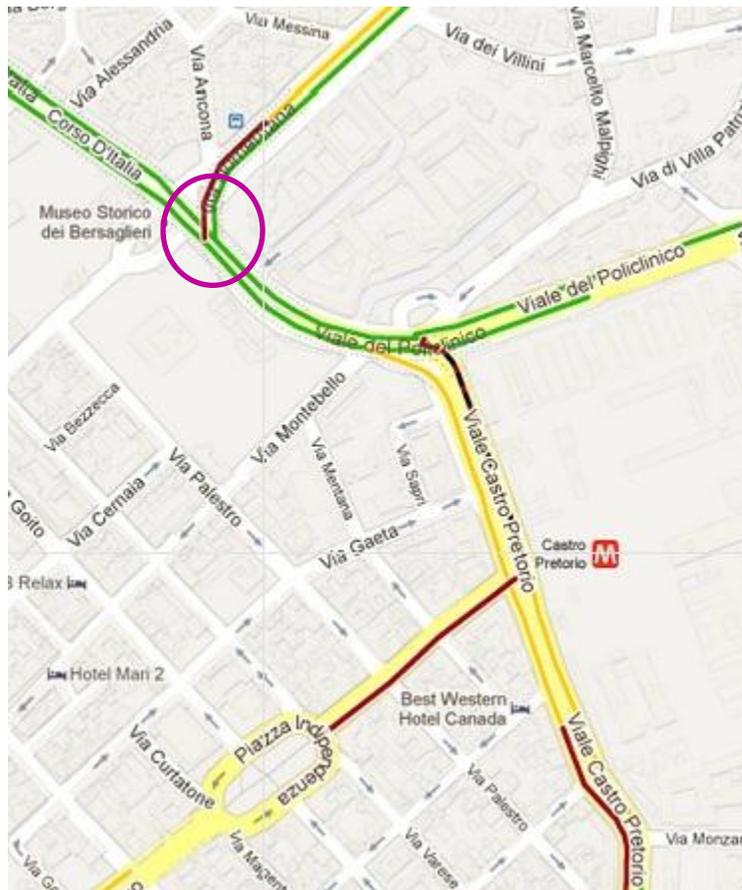
6) Repeat this process for two more of the alternative paths identified by Google maps. If no alternative paths are identified, then adjust the route on the map as described above in step 3).

7) Compare your answers to determine which of the three paths is optimal in terms of pollution cost.

Now that a minimal pollution cost can be calculated for an entire A-B path, a similar method can be developed for **solution b)**. Google maps or MapQuest can still be used for **solution b)**, however, the network path is partitioned by links (identified in the directions of the web map system) and D, TT and V are computed for each link. The distance D and travel time TT can be computed again using the same maps but with various subpaths. The subpaths are defined by the junctions (vertices) identified by the numbered written directions that appear with the map route. The value of V can be evaluated as before, assuming there are at least two different colors within each segment. A further refinement for the value of V is possible by adding a value for V_j at junctions. This added value is calculated as the average velocity difference of incident links. That is, numbering the links to the junction J from 1 to n_J, then

$$V = \frac{1}{n_J} \sum_{i=1}^{n_J-1} |Color(i+1) - Color(i)|, \tag{4}$$

(Note: n_J is the number of links to the junction J.) Then the optimization of paths over a graph is obtained by "enlarging" the graph adding links.

Example: Calculate the value of v at the junction of Corso D'Italia and Via Nomentana in Rome, Italy (identified in the diagram below).

The junction has three links: Via Nomentana (red – 15mph) and two Corso D'Italia (green – 35mph). Therefore,

$$v = \frac{1}{3}(|35 - 15| + |15 - 35|) = \frac{1}{3}(20 + 20) = \frac{40}{3}$$

Class Activity: Calculate the pollution cost (FC) for one of the paths from the previous activity (Independence Hall to Valley Forge Visitor Center) using solution b) in which the path is partitioned into subpaths.

Class Activity: Use a web mapping system to compute the "greenest" path for your commute from your home (starting location A) to your institution (destination location B). Calculate the greenest path using solution a) and solution b).

Summary

Using information available on the web and recent results from fuel consumption measurement; it is possible to determine the "greenest" path for assigned starting and destination positions. A more advanced version that partitions a path into subpaths can use optimization of paths over graphs.
Students should be able to compute the "greenest" path for their daily commutes as well as other travels.

Exercises

Exercise 1:
To have more paths than those chosen by Google maps, one can use the function "Drag to change the route" over the proposed map. Using solution a) and solution b) calculate and compare the commute from your home to your institution using at least three paths where at least one of those paths is created by dragging the mouse to change the route.

Exercise 2:
Compare the results of two different web mapping systems (for example, compare the results of Google maps with that of Map Quest) for a given map. Interpret this in terms of reliability of the information:

    a) Do the different web mapping systems give the same "primary" path?

    b) Are the calculated pollution costs the same for corresponding paths in each web mapping system (note: calculations should be done simultaneously so that traffic volume is constant)?

Exercise 3:
It is possible to use software packages such as Excel, Matlab, Maple, Mathematica, or some other computer algebra system to calculate the pollution cost for a given path. Using the software of your choice, write script that will allow you to input the critical information for $D, \; TT, \; and \; V$ and then calculate the pollution cost (FC) for several paths comparing them to determine the optimal path.

Extensions

Extension 1:

Using a local map or information from people, compute the number of traffic lights and stops for the considered paths. Correct the cost by multiplying the computed cost by the coefficient: $1 + 0.02 \cdot NTLS$, where $NTLS1$ is the number of traffic light and stops.

Extension 2:
Fix a portion of a road network and use Google maps to measure length of links and traffic load. One simple model for vehicular traffic is provided by the rule:

$$Velocity = Max_{Velocity} - Load \cdot \left(\frac{Max_{Velocity}}{Max_{Load}}\right), \qquad (1)$$

where $Max_{Velocity}$ is the maximal velocity and Max_{Load} the maximal load. Normalizing the maximum velocity as $Max_{Velocity} = 1$, then interpreting the traffic color densities (black, red, yellow, and green) as corresponding velocities, and again assuming equal distancing between the representative color velocities, we have each color velocity corresponding to: 1 for Green, 0.75 for Yellow, 0.5 for Red, and 0.25 for Black.

Exercise: Use these values to compute normalized loads.

Exercise: Assume that the pollution is directly proportional to the traffic load. Compute the polluting effect of different links or different road systems.

A further extension is possible using Reference 1 and different rules for the velocity as function of the load, known also as Fundamental Diagram.

<u>References</u>

Traffic modeling and traffic networks:

M. Garavello, B. Piccoli: Traffic flow on networks, Applied Math Series vol. 1, AmericanInstitute of Mathematical Sciences, 2006.

References for the fuel consumption model:
    http://transportation.mit.edu/live/news/1320-energy-and-emissions-logging-in-road-vehicles

    http://www.nasa.gov/topics/earth/features/health-sapping.html

    R. J. Laumbach and H. M. Kipen, "Acute effects of motor vehicle traffic-related air pollution exposures on measures of oxidative stress in human airways," *Annals of the New York Academy of Sciences*, vol. 1203, pp 107-112, August 2010