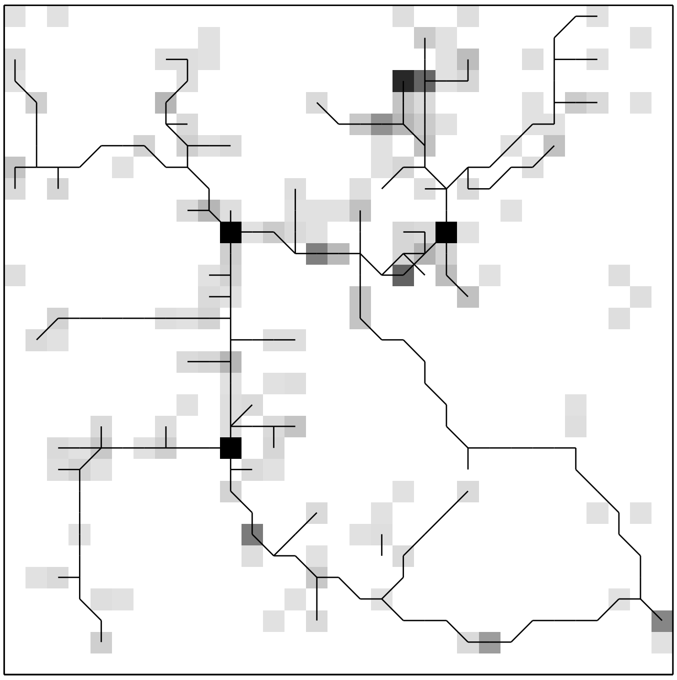
CHALMERS



Road Growth Modeling

The Development of a Road Growth Model That Co-Evolves with Land Use

Master of Science Thesis

JOHAN LODIN

Department of Energy and Environment
Division of Physical Resource Theory – Complex systems group
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden, 2011

Road Growth Modeling

The Development of a Road Growth
Model That Co-Evolves with Land Use

JOHAN LODIN

Department of Energy and Environment
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden, 2011

Road Growth Modeling
The Development of a Road Growth Model That Co-Evolves with Land Use
JOHAN LODIN

Department of Energy and Environment
Chalmers University Of Technology
SE-412 96 Göteborg
Göteborg, 2011

Cover:
The result of a simulation of the model presented in this thesis. The darker an area is the more economically active it is. The thin solid lines are roads.

Chalmers Reproservice
Göteborg, Sweden, 2011

Abstract

Land development and infrastructure growth are two co-evolving processes. When studying long-term urban evolution it is useful to model both processes in parallel but co-dependently. The focus in this thesis is a novel road growth model that extends and adapts an existing land use model. The model transforms land use into interactions—a form of abstract traffic-like quantities—between every pair land lots. Each interaction needs to find a path from its start lot to its destination. If it is more efficient to extend the existing road network then a new road is suggested. Though each land lot pair acts independently when finding the path between the pair, a level co-operation occurs in the model as suggestions may add up. The model formulation turns out to have a scale-invariant property and a possible connection to the fractal-like structures of real road networks is explored. The model is open-ended with regards to the underlying geography, but makes several simplifications; only one type of traffic is considered, all roads have infinite capacity and equal speed limits, and travel time is approximated by availability and Euclidean distance. Care has been taken to make the model implementable in a computer with a reasonable amount of working memory even when there are many land lots, at the cost of increased computing power. A large part of the model is however parallelizable and some applicable time-saving techniques are recognized. A brief overview of relevant aspects of infrastructure in general—and road networks in particular—is also presented, giving a basis for model evaluation.

Contents

1 Introduction

In urban and rural systems, how does infrastructure evolve? In particular, how does infrastructure co-develop with land use on a very long time scale? Are there some basic principles that govern the growth of infrastructure, and if so, are they the same at different length scales and in areas with different population density? Do different kinds of infrastructure have structural similarities? If we understand the principles of land use and infrastructure growth, can we then find optimal policies from *e.g.* an economical/transportation perspective?

To aid in answering the above questions simplified models may be used. For instance, traffic models have been extensively studied in operational research. The models of interest in this thesis, however, represent a continuously evolving scene where infrastructure and land use co-evolve and scales spanning a wide range are of interest both in space and time.

There is an explorative component in the project regarding both the models used and which questions that will be of interest. A part of the thesis is thus a literature review of aspects that may be of interest—or turn out to be superfluous—for modeling road growth. Another part is a road growth model co-evolving with land use which is the major result of this thesis.

The goal of this thesis is not to answer the questions above, but to hopefully to take a small step closer to answering them.

2 Method

Several aspects of the problem had to be considered simultaneously, not least the representation, relevant properties, and executability.

Implementing the model, and variations of the model, is beneficial in developing the model. Even though the results from a run of an implementation of the model may not be enough to give any quantitative results, the process helps to pin-point both theoretical and practical problems and to isolate interesting aspects. Not only does the representation become a practical issue and memory usage evident, it also demonstrates some phenomena emerging at a still fairly microscopic scale.

To find what aspects of infrastructure networks that are relevant in general, and road networks in particular, literature studies accompanied the implementation. This was an interative process where literature and model development/implementation went hand-in-hand.

The model was developed by first studying how roads would and should grow in simple systems with two or three disconnected points ("cities"). In those systems extreme cases are fairly easy to analyze and the results were generalized where applicable. Where the results were not generalizable the model was developed further.

The model was implemented in programming languages Java and Matlab. Java was used for the larger implementation. Matlab was used for several smaller setups, mini experiments, and visualizing the results from the Java implementation[1]. Java and Matlab were chosen as programming languages as they are fairly convenient and efficient to program in. The Java implementation was written for flexibility rather than execution efficiency as the primary purpose was to use it as a tool in developing the model.

3 Aspects of Infrastructure and Previous Work

To know what to model a review of interesting aspects of infrastructure is needed. The same aspects may also serve as a validation of the model, but it also highlights the lack of realism the simplification of the model implies.

3.1 Network Representation of Roads

To be able to grow and analyze an infrastructure, an encoding of the infrastructure is needed. A graph is a natural fit. A graph consists of nodes and edges that connect the nodes to each other. In a road network there are roads and intersections. The road network can thus be mapped to the graph it at least two ways. Either the roads are edges or nodes. The perhaps most intuitive is to map intersections to nodes. If there is a road directly connecting two intersections then an edge connects the nodes. In this representation any measure of the road network, such as Euclidean road length and traffic flow, is easily preserved and thus gives an objective representation of the real-world network. This representation will visually replicate a two-dimensional map but the intuitive notion of road, something that continues through several intersections, is however not immediately available. Instead the edges are better thought of as representing road segments.

The roads can instead be mapped to the graph's nodes. The edges now come to represent the intersections. This is the so called dual representation. In this representation the sense of *e.g.* Euclidean distance is lost, and one must ponder what constitutes a road. One method is to group road segments by their names, as done by Jiang and Claramunt (2004) and Kalapala et al. (2006). Another is to use geometrical properties such as line of sight or incident angles at the intersections, as done by Porta et al. (2004) and Turner (2007). For structural studies of road networks the street names are a bit arbitrary but may suffice as an approximation. Geometrical methods may on the other hand give rise to unexpected artifacts or biases but have the potential to better represent the human intuition of roads.

[1] Such a visualization is seen on the front page. The black lines are roads and the darker the underlying cell is the more active it is. The system evaluated had a fairly arbitrarily set parameters and no real conclusions can be drawn from it.

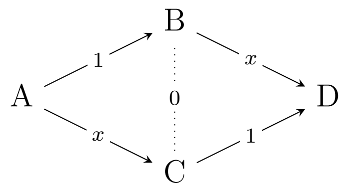


Figure 1: An example of a system in which Braess' paradox can occur when the dotted connection is added, if the total flow between A and D is between $\frac{2}{3}$ and 2. The edge labels are the travel times along the edge and x is the flow.

3.2 Behavior and Flow of Traffic

Travel time and congestion

In the 1960s the United States Federal Highway Administration used the relation

$$A_{ij}(x_{ij}) = a_{ij} + b_{ij}x_{ij}^4 \tag{1}$$

to model road travel times based on traffic flow where A_{ij} is the travel time on road ij based on the traffic flow x_{ij} (LeBlanc et al., 1975). a_{ij} and b_{ij} are road parameters. If written as

$$A_{ij}(x_{ij}) = a_{ij} + \left(\frac{x_{ij}}{b_{ij}^{-1/4}}\right)^4$$

it is clear that $b_{ij}^{-1/4}$ represents some capacity measure and congestion occurs when the flow starts to noticably exceed this value. The a_{ij} parameter represents the minimum travel time and can thus be related to the speed limit for ideal roads if $a_{ij} = l_{ij}/v_{ij}^{max}$ where l_{ij} is the road length and v_{ij}^{max} is the speed limit.

Braess' paradox

Braess' paradox is an equilibrium paradox similar to the Prisoners' Dilemma. It was published by Dietrich Braess in 1968 (Braess et al., 2005). Braess showed that user equilibrium need not minimize the travel times in the system, and that an extension of the system, *e.g.* some sort of short-cut, may make the travel times longer. The flip side of Braess' paradox is that system optimality can be achived by slowing down traffic or shutting down roads. For random graphs it has been shown that the paradox is likely to occur (Valiant and Roughgarden, 2006) and occurs in the Boston-Cambridge area (USA), London (UK), and New York City (USA) (Youn et al., 2008).

Figure 1 illustrates a topology that can exhibit Braess' paradox. Assume nodes A, B, C, and D, arranged so that there is a road between AB, BD, AC, and CD. Assume that there is a total flow of $X = 0.8$ units between A and D. The flow splits up and go by either B or C. If travel times are $t_{\mathrm{AB}}(x) = 1$, $t_{\mathrm{AC}}(x) = x$,

$t_{\mathrm{BD}}(x) = x$, and $t_{\mathrm{CD}}(x) = 1$, where x is the current flow on the road. Due to the symmetry the user equilibrium will be $x_{\mathrm{AB}} = x_{\mathrm{BD}} = x_{\mathrm{AC}} = x_{\mathrm{CD}} = X/2$. Travel time will thus be $1+X/2 = 1.4$. If a short-cut is created between B and C, so that $t_{\mathrm{BC}}(x) = 0$, then the "new" fastest route is ACBD. If only one user changes path to use the short-cut BC then its total travel time will be $X/2 + 0 + X/2 = X = 0.8$ which is less than 1.4 which is the travel time in the original system and the travel time all other users experience. Hence users will start taking the short-cut instead. If all users take route ACBD, and they will, then the travel times will be $X + X = 1.6$ which is slower than the travel time of 1.4 in the original system. Each user will choose route ACBD since nothing is gained by taking AB or CD with constant travel time 1 over AC or BD respectively with a travel time of 0.8. In this example the paradox can only occur when $2/3 < X < 2$.

The above AB-AC-BC-BD-CD structure (ignoring directivity) is the only two-terminal non-reducable undirected network topology that can exhibit Braess' paradox (Milchtaich, 2006).

Route selection

John Glen Wardrop postulated in 1952 that traffic equilibrium occurs when no user can take a different route to minimize their travel time. However, route selection is a complex task and humans try to reduce complexity to a relatively low level and navigate in a fairly straight line to the target (Dalton, 2003). Thus, microscopic modeling of users may not be as simple as finding the fastest path.

Calculating traffic flow equilibrium

The traffic assignment problem for flow equilibrium is about finding the flow on each road given a particular road network and origin-destination pairs. A bit more formally, let G be the digraph representing the road network such that the nodes are the intersections and the edges are the roads. Between each pair of nodes, (i, j), there is a directed flow of $F_{ij} \geq 0$. The "cost" of the flow x_{ij} between neighboring nodes i and j is $f_{ij}(x_{ij})$. The cost function is typically non-linear, *e.g.* (1). The flow is typically conserved, meaning that all flow out from a node comes from either a source in the node or from external flow into the node, and every source is matched by a sink. Given that the flow is conserved, the problem is to find the x_{ij} which minimizes

$$\sum_{(i,j)} f_{ij}(x_{ij}),$$

and the equilibrium found by solving this is the so called user equilibrium (LeBlanc et al., 1975). In this simplified view there are no transients or temporally varying flows or congestions. Some such effects could be added by letting the cost function $f_{ij}(x)$ depend on time and solving the problem for each time.

There are several methods for solving the traffic assignment problem. Typically these start with an initial guess, and typically the fastest useful initial guess is the solution to the shortest path problem, *i.e.*, the flow takes the shortest, not quickest, way from the origin to the destination. The solution is then refined.

The all-pairs shortest path problem is also a problem with several well-known solutions, such as Dijkstra's algorithm and the Floyd-Warshall algorithm. Dijkstra's algorithm is a greedy algorithm that visits the nearest nodes first, relative to a starting node, and thus gives the shortest path from one node to all other nodes. Dijkstra's algorithm is effective and very simple to implement. If implemented using a good choice of data structure its running time grows as $|E| + |V| \log(|V|)$ where $|E|$ is the number of edges and $|V|$ is the number of nodes. To solve the all-pairs shortest path problem Dijkstra's algorithm is performed for all nodes, giving a worse case running time that grows like $|V||E| + |V|^2 \log(|V|)$.

There is an algorithm similar to Dijkstra's called A* which finds the shortest path between two nodes. A* uses a heuristic function that, if chosen carefully, can narrow down the number of visited nodes in order to more effectively find the destination node. It is also greedy but instead of visiting nodes sorted by closeness to the start node it estimates closeness to the target node. In order to get better performance than Dijkstra's algorithm some knowledge of the network topology is needed. If the nodes are mapped to a coordinate system then the nodes' coordinates are such knowledge. If no information is provided through the heuristic function then A* is equivalent to Dijkstra's algorithm. In comparision, if the graph is a fully connected planar graph where the nodes are the cells of a square lattice, then the number of nodes that Dijkstra's algorithm has visited before finding a particular goal node grows with the square of the distance to the goal. The number of nodes that A* has visited grows linearly with the distance to the goal.

3.3 Structural Properties

Lämmer et al. (2006) did a study of the 20 largest cities in Germany. The study let the intersections be mapped to nodes in the graph and used shortest paths to approximate traffic flows. The study estimates that for Dresden 50 % of all road meters carry only 0.2 % total traffic volume. Additionally, 80 % of all the traffic is concentrated to no more than 10 % of the road meters, and only 3.2 % of the road meters handle 50 % of the traffic. These numbers, as interpreted, are a bit questionable. Total traffic is defined as the shortest distance between two nodes summed over all node pairs and the flow is evaluated using a betweenness centrality measure. It does however paint a clear picture of the structure of the road network. The study thus rather shows that only a small fraction of the road network is central, and that most of the network is of no common interest.

In the same study cell sizes and cell shapes were studied. Cell sizes were found to fit a power law with negative slope, meaning that the road network has a lot of tight loops and few large uncrossed areas. Cell shapes show a lot of diversity and the authors note that this may reflect the fact that the cities are old and therefore lack the uniformity of a planned city layout.

Interestingly, the average number of reachable intersections starting from an arbitrary intersection was found to grow polynomially as a function of travel time, barring any upper and lower cut-off effects. This is an indicator of fractality in a graph, where the exponent is the effective dimension.

A similar result was found by Kalapala et al. (2006) in which journey structure was considered in a different fashion. Large-scale road networks in USA, England, and Denmark were studied. For each sampled journey the road segments were grouped by name. The journey was then profiled by looking at the fraction of total length spent on each road. The study found that across journeys ranging from relatively short to relatively long, the longest distance spent on a road, relative to the total travel distance, was the same. The same was found for the second longest distance, and so on. The fact that the journey profile is similar for both short, medium, and long journeys shows that the road network has a fractal-like property. Porta et al. (2004) used another dual representation based on geometrical properties of the network. That study also found that the average number of nodes in a neighborhood as a function of the radius followed a power law distribution, but sampled networks from one square mile sections of six diverse cities which in contrast are very small networks. This fractal-like property is again supported by a study by Csányi and Szendrői (2004) in which it is argued that geographically constrainted networks have fractal-like scaling, giving examples from the US power grid network, London's Underground network (UK), and Hungary's water network. Further, Csányi and Szendrői (2004) argues, fractal networks and small-world networks make up a dichotomy.

However, contradicting results come from Jiang and Claramunt (2004). Their study of the street networks in Gävle (Sweden), Munich (Germany), and San Francisco (USA), using a dual representation based on names, showed the opposite. The networks had short separations and high clustering, *i.e.*, a small-world network. This naturally leads to the question about data set sensitivity as well as sensitivity to the choice of dual representation.

3.4 Construction Costs

Road construction costs can be estimated using units prices per length or area, with some adjustments made for small projects due to inefficiencies regarding *e.g.* mobilization of equipment (United States Department of Agriculture, 2009). That is, constructions costs are approximately linear with a small startup cost.

3.5 Previous Work

There exists models of both road growth and land use. Since the goal is to connect the two mechanisms, two "compatible" models can be used as a base. Next is a presentation of two models that are important to the proposed model and road growth algorithm.

3.5.1 Road Growth Models

A fractal toy model was developed by Kalapala et al. (2006), specifically designed to reproduce certain statistical properties. While that goal is achieved the resulting network cannot be considered a realistic reproduction of real road networks.

A model by Barthélemy and Flammini (2009) effectively captured important statistical properties of a city with respect to the road network, as well as actually looking like a real road network. In that model the city is mapped to a grid and an economic model is used to set up the probabilities of developing new areas (”centers” in the Barthélemy-Flammini model). Once an area is developed it is no longer an active component in the model.

The road growth mechanism in the Barthélemy-Flammini model is greedy and purely geometrical. No economic cost is placed on building a new road; it is assumed that whenever a new center is developed then new roads are built from the existing road network. All road nodes in the relative neighborhood of the new center may be starting points for new roads. In that way loops can be created.

If two new centers are created and the same piece of road is the closest point in the road network then a new small piece of road is built in the direction of a point lying between the two new centers. This is done iteratively until the road is forced to split and extend as two different road segments in opposite directions. See Barthélemy and Flammini (2009) for details. The authors note that their scheme can be generalized so that different centers have different weights, but claim that the large-scale structural properties of their model is unaffected as long as the ”heavy” and ”light” centers are uncorrelated, uniformely distributed in space, and not broad.

While the Barthélemy-Flammini model gives interesting statistical properties of the road network and looks like a road network, the growth mechanism can be questioned. In some sense the city grows from the outside in since new centers are immediately connected to the road network. There is a population density in the model, but the land use is not included in the model.

3.5.2 Economic Land Use Model

Andersson et al. (2003, 2005a,b) have developed a model of land use. After calibration with real world data the model gives the right type of distributions of cluster sizes as well land values. See Figure 2. It also gives the right dependency between cluster area and cluster perimeter length, as well as aggregated cluster land values and cluster area. Some of the model's parameters can also be estimated from measurable real-world data, making model validation easier.

The infrastructure is implicit in Andersson's model, but the model is fairly open-ended with regards to infrastructure and can thus adapted to use an explicit infrastructure model instead.

The model is situated on a regular square lattice spanning $\mathbf{R}^2$. The real plane is divided into evenly sized subsets associated with every lattice point so that every point in $\mathbf{R}^2$ can be surjectively mapped to a lattice point. There is also an injective map associating every lattice point with a point in the real plane.

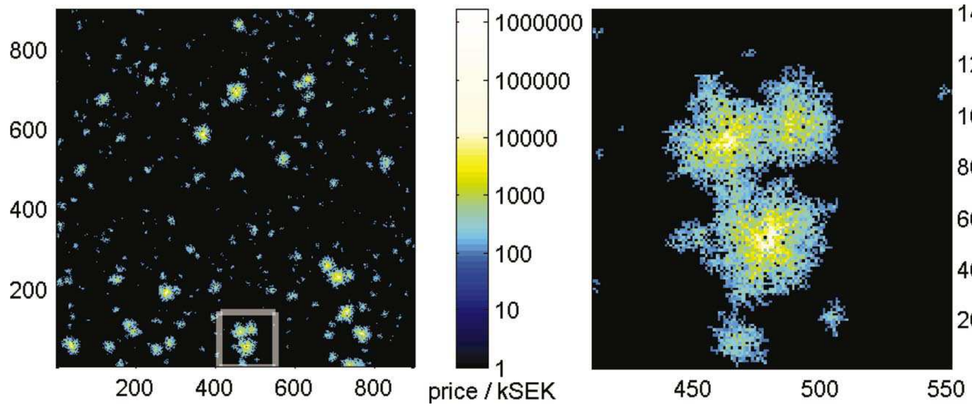


Figure 2: (Color) Simulated land prices from Andersson's model (Andersson et al., 2003). Reprinted with permission from Andersson et al., Phys. Rev. E 68, 036124 (2003). Copyright (2003) by the American Physical Society.

Each lattice point is a node in an initially completely disconnected network. The model is initialized by connecting a fixed number of nodes to at least one node in the network, including itself. All connections are undirected and the number of connections between node i and node j is x_{ij}. The degree of node i is $x_i = \sum_k x_{ik}$ and represents the economic activity in that node. The term **economic network** will refer to all cells with $x_i > 0$ (*i.e.*, the nodes) and their connections. The nodes in the economic network will be collectively denoted by the set $\mathcal{N}_{eco}$.

After initialization the iterative growth process is performed in two steps: primary node selection and secondary node selection. The primary selection is either multiplicative, with probability q, or additive, with probabiligy $(1-q)$. The probability of selecting node i during primary selection is

$$\Pi_i = \underbrace{q\frac{x_i}{\sum_k x_k}}_{\text{multiplicative}} + \underbrace{(1-q)\frac{a_i}{\sum_k a_k}}_{\text{additive}}$$

and $a_i \in [0,1]$ is the **availability** of each node. The details of how a_i is determined are subject to the infrastructure model. The secondary selection, *i.e.*, the probability that node j is selected given that node i was selected as the primary node, is

$$\Pi_{ij} = \underbrace{q\frac{a_{ij}D_{ij}x_j}{\sum_k a_{ik}D_{ik}x_k}}_{\text{multiplicative}} + \underbrace{(1-q)\frac{a_{ij}D_{ij}}{\sum_k a_{ik}D_{ik}}}_{\text{additive}} \tag{2}$$

where $a_{ij} \in [0,1]$ is the **relative availability**[2] and $D_{ij} \in [0,1]$ is the spatial interaction strength. The form suggested by Andersson et al. is

$$D_{ij} = (1 + cd_{ij})^{-\alpha} \tag{3}$$

[2]The relative availability is introduced here and is not used in the original papers. In the original papers a_j is used instead of a_{ij}. In Section 4.3 the introduction of a_{ij} is motivated.

where α and c are system parameters and d_{ij} is the Euclidean distance. Both D_{ij} and a_{ij} should be symmetric.

When the primary and secondary nodes, i and j, have been selected then a connection between them is added, *i.e.*, x_{ij} is increased by one.

The additive growth is not related to x_i. That means that nodes with $x_i = 0$ may be selected. This is the process that extends the economic network to more nodes. It is the combination of introduction of new nodes and the multiplicative growth that gives rise to a power-law distribution of the node degrees (Barabasi, 1999).

The spatiality of the model enters via the availability a_i, relative availability a_{ij}, and the interaction strength D_{ij}. The availability may depend on the proximity of infrastructure, relative availability may depend on the transportation possibilities between the nodes, and the interaction strength may depend on *e.g.* total travel time between the nodes.

The original model, as formulated by Andersson et al., has only a very simple ambient infrastructure model where a_i is directly related to the activity, x_i. If $x_i > 0$ then $a_i = 1$, *i.e.*, full availability. If $x_i = 0$ and a neighbor i' to i has $x_{i'} > 0$ then $a_i = a^{(\mathrm{P})}$, *i.e.*, "perimeter availability". If $x_i = 0$ and i does not have any active neighbor, then $a_i = a^{(\mathrm{E})}$, *i.e.*, "external availability".

4 Connecting Land Use Growth and Road Growth

The Barthélemy-Flammini model and Andersson's model make an interesting couple. While the infrastructure is implicitly present in Andersson's model via the availability classification of cells, the infrastructure and its centrality is explicitly present in the Barthélemy-Flammini model. On the other hand, while the Barthélemy-Flammini model spawns fully active centers which becomes passive once connected to the road network, Andersson's model has increasingly active centers which never become passive.

The model presented below is a novel model that extends Andersson's model with a road growth mechnism inspired by the Barthélemy-Flammini model. It is a planar model driven by the needs to interact, and roads are built where it is considered economically beneficial.

The model is iteratively evaluated in much the same way as Andersson's model. Between "economic" iterations—when the activity is increased—the road growth algorithm is evaluated and roads may be constructed. When the road growth algorithm is completed another economic iteration may occur.

As in Andersson's model cells and nodes are overlapping terms. Cells and nodes will be used almost interchangeably throughout the text and the context will determine which term is used. The terms road network and infrastructure will also be used interchangeably, accompanied with the terms traffic and interactions.

4.1 Assumptions and Limitations

In a planar model all roads that cross each other intersect. The model can thus not represent viaducts or air travel, for instance. Individual roads are not modeled *per se*, but rather traffic connections between different areas. Despite this the connections will be referred to as roads and road segments. Each connection is assumed to have enough capacity to carry all the traffic that may use it.

When the economy grows the spatial impact is assumed to be low. In particular, the discrepancy between the actual travel distance and the Euclidean distance between two cells is assumed to be negligible. The relative availability however is assumed to dominate the selection.

The road growth process is assumed to be a slow process compared to trade evolution, *i.e.*, beneficial extensions to the road network is immediately taken advantage of and trade adapts immediately.

4.2 Model Representation

The infrastructure has a representation similar to the economic activity. The infrastructure is a network consisting of a set of nodes, $\mathcal{N}_{IS}$, and a set of connections, $\mathcal{R}$, and a spatial association. There is a surjective map $\mathbf{R}^2 \to \mathcal{N}_{IS}$ mapping all points in the real plane to the corresponding infrastructure node. All points mapped to the same node define an infrastructure cell. The infrastructure cells need not be of the same size as the economic cells. There is also an injective map $Coord_{IS} : \mathcal{N}_{IS} \to \mathbf{R}^2$ giving a coordinate for all infrastructure nodes.

The **reach** of an infrastructure node in the economic network is determined by first mapping the economic node to a point in the real plane, and then mapping that point to a node in the infrastructure network. All economic nodes that map to an infrastructure node make up the reach of the infrastructure node. If the infrastructure cells are smaller than the economic cells then some infrastructure cells will have no reach.

A **road**, *i.e.*, an edge $e_{ij} : e_{ij} \in \mathcal{R}, i, j \in \mathcal{N}_{IS}$, can only connect a cell to itself or to a neighboring cell. Initially $\mathcal{R} = \emptyset$. Roads between cells are added through Algorithm 1 in Section 4.4. Additionally, all infrastructure cells that have any economic activity within its reach have internal infrastructure, *i.e.*, a "road" to itself. All infrastructure nodes that have at least one edge have an edge to itself. That is of no great importance but simplifies the notion of "having infrastructure", *i.e.*, the node u has infrastructure if $e_{uu} \in \mathcal{R}$.

Since the infrastructure nodes define the intersections two edges cannot intersect spatially. Therefore some neighbors in the spatial sense must be excluded as network neighbors when the network grows. For instance, if the lattice is square and the neighborhood is a Moore neighborhood, then if the north and east cells are connected, then the north-east cell is not a neighbor to the center cell anymore.

4.3 Adjustments to the Land Use Model

The road growth model is naturally driven by the activity in the land use model. At the same time the roads can affect the land use model. The original land use model has two measures that reflect the infrastructure; the availability and the spatial interaction strength. In the land use model there is an implicit assumption that there is some form of background infrastructure making all nodes reachable. These two measures need to be adjusted to take into account the explicit infrastructure.

The most obvious change with regard to availability is the fact that in the presence of roads there exist cells with no activity but with full availability. The natural extension of the availability in the primary selection is thus that infrastructure nodes that have infrastructure have full availability, and cells in the economic network have the same availability as the infrastructure node in which reach the cells lie. Nodes that do not have any infrastructure but have at least one neighbor that has infrastructure have perimeter availability, and this availability is again propagated down to the underlying economic network.

Another consideration regarding the availability is how disconnected infrastructure components should be treated in the secondary selection. A simple addition to the original model is a **relative availability**[3], a_{ij}. The relative availability is defined as the minimum availability one must to pass to go from cell i to cell j, see Figure 3, with the special case $a_{ii} = a_i$. That is, all nodes reachable via infrastructure have availability $a_{ij} = 1$. All other nodes have a maximum availability of a perimeter cell, $a^{(\mathrm{P})}$. If a perimeter cell but no external cell is needed to be passed to get between two infrastructure components then all cells in the other component are perimeter cells. All nodes that are not reachable via infrastructure or via perimeter cells are external cells and have availability $a^{(\mathrm{E})}$. For instance, a cluster of neighboring nodes having only internal infrastructure all have a relative infrastructure of $a^{(\mathrm{P})}$ between each other. The relative availability as defined here expresses the assumption that passing a single perimeter or external cell dominates the availability.

In the original formulation of the land use model the spatial interaction strength D_{ij} is on the form

$$D_{ij} = (1 + cd_{ij})^{-\alpha}.$$

Instead of using the Euclidean distance d_{ij} the travel time between i and j could be used. Nodes that are not reachable through the road network need then be considered. However, this is in a way already handled by the relative availability, and for simplicity's sake D_{ij} is kept on its original form.

4.4 Road Growth Algorithm

The core of the road construction algorithm is given in Algorithm 1. The first unexplained entity is the interaction measure b_{ij}. The economic network is mapped to interactions in the infrastructure network. The interactions can be mapped in

[3]The relative availability was already introduced in (2).

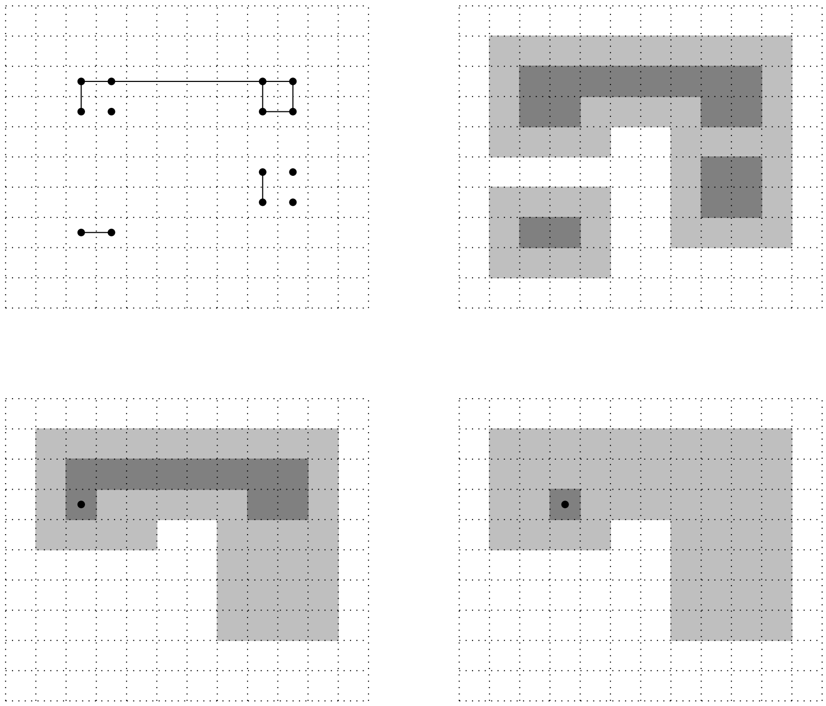


Figure 3: Upper left: a simple system. The dots mark active cells and the solid lines are roads. Upper right: The availability, a_i, used in the primary cell selection. Dark gray means full availability, light gray means perimeter availability, $a^{(\mathrm{P})}$, and white cells are external, $a^{(\mathrm{E})}$. Lower left: The relative availability, a_{ij}, used in the secondary selection, as seen from the cell marked with a dot. Lower right: The relative availability from another cell. As seen, the roads affect both a_i and a_{ij}. The road connecting the two upper clusters of active cells makes the four inactive cells in between fully available. In the absense of roads the relative availability instead isolates disconnected components, for instance the two connected cells in the lower left part of the system.

several ways which will be discussed later. Interactions are not real traffic flows but rather the means of evaluating where roads are needed. The interactions need to find a path between the origin node and the destination node. The cost of traversing the infrastructure network via roads—the "transportation" cost—is proportional to the road length and the size of the interaction. The system parameter $\hat{c}^{(\mathrm{T})}$ is the cost for an interaction unit to travel one distance unit in the road network. The cost of going off-road is $\hat{c}^{(\mathrm{T})}l^{(\mathrm{E})}$ per unit distance, and the unit cost of constructing a road segment is expressed in system parameter $\hat{c}^{(\mathrm{C})}$. $l^{(\mathrm{E})}$ is the system parameter controlling the relative cost increase for traveling externally to the road network, *i.e.*, going off-road.

The $W(c_{\mathrm{on-road}}, c_{\mathrm{off-road}})$ function takes two cost arguments and returns an adjacency matrix connecting the infrastructure nodes. It is defined by

$$(W(w_{\mathrm{on-road}}, w_{\mathrm{off-road}}))_{uv} = \begin{cases} \infty & \text{if } u \text{ and } v \text{ are not neighbors} \\ w_{\mathrm{on-road}} d_{uv} & \text{if } e_{uv} \in \mathcal{R} \\ w_{\mathrm{off-road}} d_{uv} & \text{if } e_{uv} \notin \mathcal{R} \end{cases}$$

where d_{uv} is the Euclidean distance between nodes u and v.

The shortest path between two nodes using the edge weights $\underset{\sim}{w}$ is denoted by $\mathrm{Path}(\underset{\sim}{w}; u, v)$. The path length of path P given a weight matrix $\underset{\sim}{w}$, *i.e.*, the sum of the edge weights for all edges in the path, is $L(\underset{\sim}{w}; P)$. So $L(\underset{\sim}{w}^{(\mathrm{C})}; P')$ in Algorithm 1 is the "construction" cost along path P' and $L(\underset{\sim}{w}^{(\mathrm{T})}; P')$ is the transportation cost.

The $\underset{\sim}{w}$ matrix in Algorithm 1 represents the transportation costs in the current road network. In the $\underset{\sim}{w}'$ matrix all transportation costs are the same, but if there is no road for a particular edge in the current road network then an additional cost—a "construction" cost—is added. The construction cost should not be understood as an actual construction cost. It is rather the relation between the "transportation" cost and the "construction" cost that is of relevance. This will be discussed later.

If P' goes off-road then a road construction request is added at the first node after which P' goes off-road. The request is added such that it points in the direction of the first following node that is taking the route on-road, or the last node if it never goes on-road again. Requests are added up in the "request vector" $\mathbf{r}_u$ which is some sort of compromise of all requests at node u. The $\mathbf{r}_u$ vector is made up by $b_{ij}^{(\mathrm{real})}/b_{ij}^{*}$ long pieces pointing in different directions. $b_{ij}^{(\mathrm{real})}$ is the real transportation need between i and j and b_{ij}^{*} is the threshold for which the proposed new road would be economically motivated. b_{ij}, b_{ij}^{*}, and $\mathbf{r}_u$ will be discussed later. The "real" interactions, $b_{ij}^{(\mathrm{real})}$, are related to the economic network. An easy way to set up the interactions is to make use of the same statistics that govern economic growth and let

$$b_{ij}^{(\mathrm{real})} = x_i \frac{a_{ij} D_{ij} x_j}{\sum_k a_{ik} D_{ik} x_k}. \tag{4}$$

This is the multiplicative part of the secondary selection probability (2), normalized to the activity in the cell. That is, all activity in the cell is distributed over all cells in the network, proportionally to the availability, interaction strength, and activity.

Algorithm 1 The core of the road construction algorithm.

```
changed ← true
while changed is true do
  changed ← false
  for each node u ∈ N_IS do
    r_u ← 0
  end for
  for each tuple (i,j) ∈ N_IS × N_IS do
    Let b_ij be the navigational interaction measure between i and j.
    if b_ij > b_ij,max then
      b_ij ← b_ij,max
    end if
    w = ĉ^(T) b_ij W(1, l^(E))
    w^(T) = ĉ^(T) b_ij W(1, 1)
    w^(C) = ĉ^(C) W(0, 1)
    w' = w^(T) + w^(C)
    P' = Path(w'; i,j)
    if ∃e_uv ∈ P' : e_uv ∉ R then
      P = Path(w; i,j)
      b*_ij = b_ij L(w^(C);P') / (L(w;P) − L(w^(T);P'))
      Let i' and j' be the first and last nodes in the first off-road sub-path in P'.
      d = Coord_IS(j') − Coord_IS(i')
      Let b_ij^(real) be the real interaction measure between i and j.
      r_i' ← r_i' + (b_ij^(real) / b*_ij) d/|d|
    end if
  end for
  for each node u ∈ N_IS do
    if |r_u| ≥ 1 then
      Let v be the neighbor of u that lies in the direction of r_u.
      R ← R ∪ {e_uu, e_uv, e_vv}
      changed ← true
    end if
  end for
end while
```

Where (in LaTeX): $\underset{\sim}{w} = \hat{c}^{(\mathrm{T})} b_{ij} W(1, l^{(\mathrm{E})})$, $\underset{\sim}{w}^{(\mathrm{T})} = \hat{c}^{(\mathrm{T})} b_{ij} W(1,1)$, $\underset{\sim}{w}^{(\mathrm{C})} = \hat{c}^{(\mathrm{C})} W(0,1)$, $\underset{\sim}{w}' = \underset{\sim}{w}^{(\mathrm{T})} + \underset{\sim}{w}^{(\mathrm{C})}$, $P' = \mathrm{Path}(\underset{\sim}{w}'; i,j)$, $\exists e_{uv} \in P' : e_{uv} \notin \mathcal{R}$, $b^*_{ij} = b_{ij} \frac{L(\underset{\sim}{w}^{(\mathrm{C})};P')}{L(\underset{\sim}{w};P) - L(\underset{\sim}{w}^{(\mathrm{T})};P')}$, $\mathbf{d} = Coord_{IS}(j') - Coord_{IS}(i')$, $\mathbf{r}_{i'} \leftarrow \mathbf{r}_{i'} + \frac{b_{ij}^{(\mathrm{real})}}{b^*_{ij}} \frac{\mathbf{d}}{|\mathbf{d}|}$, $\mathcal{R} \leftarrow \mathcal{R} \cup \{e_{uu}, e_{uv}, e_{vv}\}$.

5 Model Analysis

Some aspects of the model can be analytically derived from the model formulation. In combination they can be the basis for less rigorously derived but more large-scale qualitative aspects of the model.

5.1 Interaction Thresholds

The transportation and construction costs determine the shortest path in the construction graph. The larger the construction cost the more does the interaction seek to use the existing road network. The shortest path in the construction graph is the path P' that minimizes

$$l_2 b_{ij} \hat{c}^{(\mathrm{T})} + l_C \hat{c}^{(\mathrm{C})} \tag{5}$$

where b_{ij} is the interaction measure between i and j, $l_2 = L(W(1,1); P')$ is the total distance when taking the shortest (partly) off-road route, and $l_C = L(W(0,1); P')$ is the total off-road distance. To construct a new road (5) must be smaller than the transportation cost in the current road network,

$$l_1 b_{ij} \hat{c}^{(\mathrm{T})} > l_2 b_{ij} \hat{c}^{(\mathrm{T})} + l_C \hat{c}^{(\mathrm{C})}, \tag{6}$$

where $l_1 = L(W(1, l^{(\mathrm{E})}), P)$ is the total distance via shortest path P between i and j in the current road network. For convenience, when calculating l_1, off-road sub-paths are considered $l^{(\mathrm{E})}$ as long as on-road sub-paths, instead of the unit cost being greater. (When calculating l_2 the whole path is considered on-road since the construction cost is added.) By necessity $l_1 \geq l_2 \geq l_C$.

Let $\Delta l = l_1 - l_2$ be the decrease in distance gained by constructing the in total l_C long road segments. The inequality (6) can then be rewritten as

$$b_{ij} > \frac{l_C}{\Delta l} \frac{\hat{c}^{(\mathrm{C})}}{\hat{c}^{(\mathrm{T})}} \mathrel{\widehat{=}} b^*_{ij} \tag{7}$$

and the threshold is denoted by b^*_{ij}. The right-hand side can be read as a product of two fractions rather than one fraction. The first fraction is $l_C/\Delta l$, which is the inverse of the relative gain a road construction would give. The second is $\hat{c}^{(\mathrm{C})}/\hat{c}^{(\mathrm{T})}$, which is a system parameter. The inequality can instead be written as

$$\frac{\Delta l}{l_C} > \frac{1}{b_{ij}} \frac{\hat{c}^{(\mathrm{C})}}{\hat{c}^{(\mathrm{T})}}.$$

Here is it clear that $\hat{c}^{(\mathrm{C})}/\hat{c}^{(\mathrm{T})}$ is the parameter that controls how much gain is needed in order to suggest a new road given a particular interaction b_{ij}. However,

$$\min \frac{\Delta l}{l_C} \to 0 \text{ when } b_{ij} \to \infty$$

i.e., if there is a straighter route between i and j then it will be built no matter what the cost when b_{ij} is large enough. Effectively a new parallel road can be

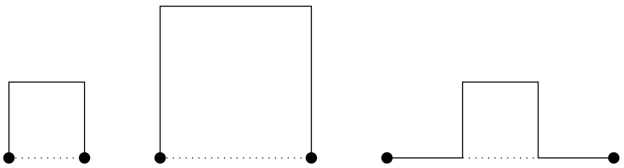


Figure 4: Three simple systems. Exactly the same amount of interaction between the dots is needed in all three systems to trigger road growth covering the dotted paths. $\Delta l/l_C = 2$ for the dotted paths. Once growth is triggered Δl will remain unchanged but l_C will decrease resulting in all of the road being constructed without any need to increase the interaction.

constructed just to avoid going one cell perpendicularly to the destination. To avoid this a minimum gain can be required yielding

$$\frac{1}{b_{ij}}\frac{\hat{c}^{(\mathrm{C})}}{\hat{c}^{(\mathrm{T})}} > \left(\frac{\Delta l}{l_C}\right)_{min} \qquad \Rightarrow \qquad b_{ij} < \frac{\hat{c}^{(\mathrm{C})}/\hat{c}^{(\mathrm{T})}}{(\Delta l/l_C)_{min}} \mathrel{\widehat{=}} b_{ij,max}$$

This limitation should only be used when setting up the $\underset{\sim}{w}'$ graph.

One could put a limitation on $\Delta l/l_C$ instead, and simply not add a request at i' if $\Delta l/l_C$ is not large enough. However, that would either potentially miss another useful request, or require a set of new iterations where the previously proposed (i',j') pairs would be excluded somehow in the path finding algorithm. By setting the limit on b_{ij} no adjustments to the algorithm are needed.

An implication of $\Delta l/l_C$ in (7) is that the threshold b^*_{ij} is scale invariant. This is a direct consequence of that all costs are depending linearly on the distance. The effect of Δl is that short-cuts are evaluated "locally" and the distance of getting to the short-cut cancels out in (6). See Figure 4.

Conceptually there is no difference between creating a short-cut and adding a new road between active cells which previously were disconnected in the road network.

A special case is a cell which has $a_{ij} \neq 1$ to all other cells. The interaction threshold (7) is then

$$b^*_{ij} = \frac{1}{l^{(\mathrm{E})} - 1}\frac{\hat{c}^{(\mathrm{C})}}{\hat{c}^{(\mathrm{T})}} \mathrel{\hat{=}} b^*_{\mathrm{Ext}}. \tag{8}$$

Notably this is independent of (i,j). It depends only on system parameters and in particular not on the distance to the nearest road network.

5.2 Single-Cell Systems

If there are only two cells, i and j, and $b_{ij} = b^{(\mathrm{real})}_{ij}$, then the interactions are

$$b_{ii} = x_i \frac{x_i}{x_i + a_{ij}D_{ij}x_j},$$
$$b_{ij} = x_i \frac{a_{ij}D_{ij}x_j}{x_i + a_{ij}D_{ij}x_j},$$

and similarly for b_{ji} and b_{jj}. The extreme case of one cell dominating the other in terms of activity reduces to

$$\lim_{x_i \to \infty} \begin{pmatrix} b_{ii} & b_{ij} \\ b_{ji} & b_{jj} \end{pmatrix} = \begin{pmatrix} \infty & a_{ij} D_{ij} x_j \\ x_j & 0 \end{pmatrix} \tag{9}$$

i.e., when one cell completely dominates then only a maximum of $a_{ij} D_{ij} \leq 1$ of the other cell's activity leaves the dominating cell, and all of the other cell's activity leaves to the dominating cell. This also means that road growth is more likely to be triggered in active perimeter cells rather than at the road network that the locally dominant cell is in.

Requests at the same road node in the same direction add up. For instance, if two requests point in the same direction they only need to be half of their respective thresholds for the road to be constructed. Requests that point in opposite directions cancel, so two requests of the same size in opposite directions will never trigger a road growth no matter the size of the requests. This is illustrated by a system with three cells that lie in a straight line with no roads between them. The threshold for road growth, b^*_{Ext}, is the same everywhere, and if all cells interact by the same amount b then road growth will start at the outer cells when $b = b^*_{\text{Ext}}/2$. The middle cell will never manage to trigger road growth.

A slightly more complicated system is if three cells are placed in an isosceles triangle where one angle is 2α and the other two angles are β, oriented so that the two edges with angles β lie on the x axis, denoted by B and B′, and the third cell on the y axis, denoted by A. See Figure 5(a). There are no roads so the thresholds are the same everywhere, b^*_{Ext}, and all cells interact by the same amount b. The request at the cell with angle 2α is

$$\begin{aligned} \mathbf{r}_{\text{A}} &= \frac{b}{b^*_{\text{Ext}}} \overrightarrow{AB} + \frac{b}{b^*_{\text{Ext}}} \overrightarrow{AB'} \\ &= \frac{b}{b^*_{\text{Ext}}} (\sin(\alpha), -\cos(\alpha)) + \frac{b}{b^*_{\text{Ext}}} (-\sin(\alpha), -\cos(\alpha)) \\ &= 2\cos(\alpha) \frac{b}{b^*_{\text{Ext}}} (0, -1). \end{aligned}$$

Road growth is triggered when $|\mathbf{r}_{\text{A}}| \geq 1$, giving

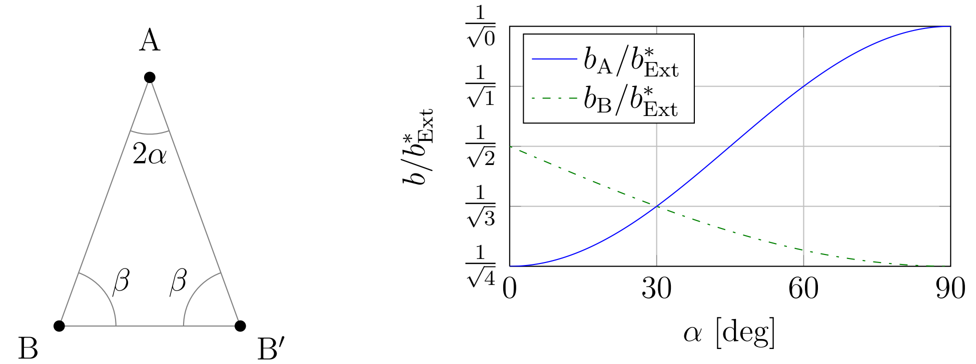
$$\frac{b}{b^*_{\text{Ext}}} \geq \frac{1}{2\cos(\alpha)} \mathrel{\widehat{=}} \frac{b_{\text{A}}}{b^*_{\text{Ext}}}$$

for node A.

The corresponding calculation for the other two nodes gives

$$\frac{b}{b^*_{\text{Ext}}} \geq \frac{1}{\sqrt{2\cos(90^\circ - \alpha) + 2}} \mathrel{\widehat{=}} \frac{b_{\text{B}}}{b^*_{\text{Ext}}}.$$

The critical amount of interaction for the A and B nodes are plotted in Figure 5(b). First we see that in the extreme where $\alpha = 0$, only half of the threshold is needed



(a) The geometry of the system.

(b) The critital amount of interaction needed to trigger road growth at node A and B, relative to the amount needed if there were only two nodes in the system.

Figure 5: A simple isoceles triangle system with no roads. All node pairs are assumed to interact by the same amount b. b^*_{Ext} is the amount that would be needed to trigger road growth at either node if there were only two nodes and no roads. Since node A interacts with both B and B′ by the same amount b then the required b to trigger road growth at A may be as low as half of b^*_{Ext} in one extreme and infinite in the other. See the text for a derivation of the functions in (b).

since the two requests add up. In the other extreme we have the situation of all cells being on a straight line, and a singularity occurs for the middle node while the other two nodes need only half of the threshold since the requests again add up.

From Figure 5(b) we can read out how the road growth will progress if b increases smoothly and $\alpha \approx 0$. At $b = b^*_{\text{Ext}}/2$ road growth is triggered at node A since the two requests add up. When the road comes closer to B the direction of the two requests start to diverge, *i.e.*, α increases, requiring larger b. When b is large enough some more road is grown. However it is not until $\alpha = 30°$ that growth also occurs from the B nodes. This is exactly when the growth nodes form an equilateral triangle and the B nodes grows in an angle of 30° to the x axis. This is the optimal geometry to minimize the road length[4]. After growth the nodes form a smaller but still equilateral triangle and this continues until the roads meet.

However, if the start condition is *e.g.* $\alpha = 60°$ then the resulting road network will not be optimal. The B nodes will trigger growth long before any road is grown at A. The B nodes will not grow roads directly towards A, which would be the optimal solution. The roads will instead intersect a bit below A.

It is very unlikely that growth will be triggered at exactly the same time at B and B' in a run of an implementation of the model. Several other scenarios are possible. If A is dominant and B is more active than B', then growth will initially

[4]This is realized by first finding that the optimal way to connect three equidistant points is to connect all of them to the centroid. If a point then is moved directly away from the centroid, the shortest additional road to add is exactly the path that the point moved, and *vice versa*.

(a) The two clusters are disconnected and aggregation of requests naturally occurs in the nodes closest to the other cluster.

(b) The clusters are connected with a road and aggregation of requests no longer naturally occurs without proper use of b_{ij}.

Figure 6: Two systems that should be equivalent, given that $l^{(\mathrm{E})}$ is set accordingly. To achive this an aggregation of interactions into b_{ij} is needed.

be triggered in A and go more towards B. Eventually growth will be triggered at B in the direction of the road connected to A but slightly drawn towards B'. When the roads from A and B meet the interactions are remapped since the availability a_{AB} change. A and B will increase their interaction at the cost of interactions to B'. The interaction from B' will however remain the same. Depending on the new distribution of interactions, A and B may add up their requests in the point on the road closest to B' and trigger road growth, or road growth may be triggered and B'.

Alternatively, if A is not active enough, B and B' will connect first. As road grows from B' to B, the requests in A will get closer and closer in direction—one pointing to the end of the road and the other at B—and possibly trigger road growth. Alternatively, when B and B' are connected there will be a point in the road network which is closest to A at which both B and B' will place their requests in the direction of A. Depending on the activity distribution, growth may occur either at A or at the road connecting B and B'.

5.3 Cluster Systems

As active cells grow roads between them, disconnected road components start to form. A road component consists of all cells that have $a_{ij} = 1$ to each other. A single cell with no external roads is thus a one-cell component.

If two components are fairly compact and well connected, then all interaction from one component to the other will end up in the road node closest to any other road node in the other component. See Figure 6(a). Since the clusters are disconnected, requests will be placed even though the individual interactions are well below the interaction threshold (8). The requests will therefore constructively add up in the closest node, and road growth will be triggered when the sum of all interactions from one component to the other reaches b^*_{Ext}.

If the isosceles triangle system above would not consist of individual cells but rather three clusters then the accumulation for large α would probably not occur. It would only occur if there is a node that is the closest one to both the other clusters, which is more likely when α is small. Neither for large α would accumulation be likely in the B nodes. The interactions going from *e.g.* B to A and B' would result in two requests in different nodes. However, if road growth is triggered anywhere in A then the new road will hold the point closest to both B nodes if α is small enough. See Figure 7.

If the active cells in Figure 6(a) remain the same but a road connects the two components, making it one big component, see Figure 6(b), then the behavior changes qualitatively. Assume that $l^{(\mathrm{E})}$ is set so that the two dotted paths in Figure 6 have the same interaction threshold. The situation is essentially the same yet road growth will not be triggered until there is at least one interaction between the clusters that is as large or larger than the interaction threshold. This is because when the interactions are smaller than the threshold they follow the road connecting the two clusters. Therefore no requests are placed.

Qualitatively nothing has changed between the two systems. They have the same activities and the same threshold for constructing a new road. It is only an artefact of the road growth algorithm that causes the change in behavior. A similar artifact arises when the resolution of the trade cells change. If the resolution is decreased, *e.g.* so that all active cells would lie in only two larger cells, then there would be no clusters and the interactions would be accumulated in the new cells, eliminating the need for constructive addition of requests. The resolution of the model thus qualitatively decide the behavior.

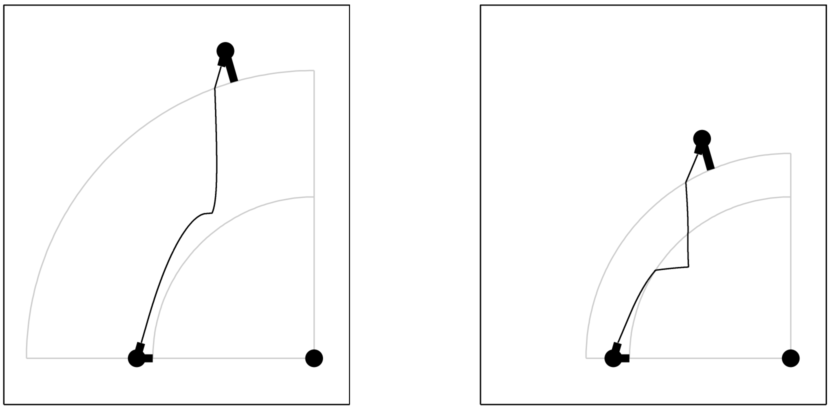
The solution is to let b_{ij} be an aggregation of real interactions. In the example above, if b_{ij} is the sum of all interactions $b_{ij}^{(\mathrm{real})}$ going from the cluster that i is in to the cluster that j is in, then requests will be placed at the dotted path. The two systems now behave the same. How b_{ij} should relate to $b_{ij}^{(\mathrm{real})}$ in general is not obvious.

If a system such as that in Figure 6(b) is scaled then the interaction threshold for road growth is constant. The interactions $b_{ij}^{(\mathrm{real})}$ are however affected. By increasing the distance between the clusters the interactions between the clusters decrease due to decreasing interaction strength D_{ij}. If the resolution is decreased proportionally to the scaling of the system natural aggregation occurs. Nearby cells aggregate into a larger single cell. The bigger the scale, the more aggregation occurs. b_{ij} must thus scale as well to keep the high and low resolution systems equivalent. How many of the neighboring nodes of i and j are included in b_{ij} should thus be a function of the distance between i and j.

In another view, b_{ij} expresses the idea that if everyone in a neighborhood agrees to contribute to the construction of a new road proportional to how much they will use it, then a more expensive road can be suggested.

6 Computational Considerations

The model is computationally heavy. To reduce the computational burden some information can be stored in the work memory. However, care must be taken to not exceed the reasonable amount of work memory available to an implementation of the model.



(a) (b)

Figure 7: The isosceles triangle system (Figure 5(a)) revisited. The three dots now represent clusters so that no road in a cluster is closest to both of the other clusters. The thick lines are existing roads. The thin line is the new road and the light gray quarter circles are centered around the bottom right dot. The new road is the result of a Matlab simulation where the interactions between the top and bottom left clusters are precisely enough to trigger road growth at both ends and all other interactions are just slightly below the threshold. Since the angle between the roads at the top cluster is small the new road quickly becomes the closest point to the bottom right cluster. Both interactions from the top cluster to the bottom clusters then add up in the end of the new road, and the road makes a sharp turn. Consequently the road from the bottom left cluster also starts to turn. In (a) the new road does not turn enough so the closest road to the bottom right cluster is still the road at the bottom left cluster. In (b) the top cluster is moved a bit closer, effectively decreasing the angle at the bottom left cluster, resulting in a sharp turn in the road from the bottom cluster as well. The twists and turns of the new roads in these examples are greatly exaggerated since the interactions to/from the bottom right cluster are just below the threshold.

For instance, if a system is modeled on a square lattice with $N = L^2$ cells, and a double precision floating point number, taking about 8 bytes, is assigned to each pair of nodes, then the total amount of memory needed for a matrix holding that information is approximately

$$M \approx 8N^2 = 8L^4.$$

If a maximum of *e.g.* two gigabytes of memory is available for the matrix, then L cannot be larger than about 128. Seeing that the memory grows as a power of four of the lattice width, any quantity relating each node to every other node is not practically storable in the working memory of a regular computer if large lattices are of interest. a_{ij} and $b_{ij}^{(\mathrm{real})}$ are examples of such quantities.

In the next sections two memory-saving techniques and two time-saving techniques are described. Three of the four techniques are connected. The first is how to efficiently maintain information about the relative availability. This does not only result in fast calculations of the relative availability but it also gives a way to efficiently maintain information about the interactions. Using the readily available interactions a time-saving technique is recognized, and finally another independent time-saving technique is noted.

6.1 Calculating Relative Availability

Since the relative availability cannot be stored in a direct form in the working memory of a regular computer it has to be calculated when needed. Calculating the relative availability by following the road network for every node pair is an expensive operation. Since the road network only grows—no nodes or edges are ever removed—it is easy to iteratively keep track of road components. A table with $O(1)$ access can be used to maintain which road component a cell is in. The same or a similar scheme can be used to maintain information about perimeter cells. When either a perimeter component or a road component is joined then the smaller component is rewritten to belong to the larger component. By comparing the values in the tables the relative availability can immediately be found, at the cost of L^2 memory compared to L^4.

6.2 Calculating Interaction Normalization

The denominator in (4), $\sum_k a_{ik} D_{ik} x_k$, is another expensive operation as it is a sum over all cells. The sum is however independent of j. The normalization can thus be maintained for all cells in only L^2 memory. Initially normalization for cells with zero activity can be ignored. When x_i first becomes greater than zero the normalization is calculated. When x_i is increased by Δx_i then normalization is increased by $a_{ki} D_{ki} \Delta x_i$ for all nodes k. Similarly, when the availability changes normalization needs to be adjusted by $\Delta a_{ij} D_{ij} x_i$ for all affected nodes i and j. The affected cells can be found by using the tables described in Section 6.1.

6.3 Stopping Short

If a cell is external to everything then road growth at that cell can not occur while

$$\sum_{k \neq i} b_{ik}^{(\mathrm{real})} < b_{\mathrm{Ext}}^{*}. \tag{10}$$

This means that the road algorithm can be skipped for all i for which (10) is true, which potentially could be a substantial computational saving. The sum over all nodes k can be avoided since all interaction sums up to the activity in the cell,

$$\sum_{k \neq i} b_{ik}^{(\mathrm{real})} = x_i - b_{ii}^{(\mathrm{real})},$$

and $b_{ii}^{(\mathrm{real})}$ is easy to compute since the denominator in the expression for $b_{ii}^{(\mathrm{real})}$ is cached as described in Section 6.2.

A generalization to (10) is to handle not only one-cell road components but also road components connecting several cells. This can make for a great computational saving since one can skip some (i,j) pairs in Algorithm 1. When the total amount of out-going interaction is less than what is needed to make any request vector in the component reach unity, then all the remaining (i,j) pairs where j is external to the component can be skipped.

The initial sum of all out-going interaction in a component $\mathcal{C}$

$$b_{\mathrm{Ext}}^{\mathcal{C}} = \sum_{i \in \mathcal{C}} x_i - \sum_{i,j \in \mathcal{C}} b_{ij}.$$

The component $\mathcal{C}$ is again easy to find using the tables used for calculating relative availability. As (i,j) pairs are being looped over and requests are being placed, $b_{\mathrm{Ext}}^{\mathcal{C}}$ is decreased. When

$$\frac{b_{\mathrm{Ext}}^{\mathcal{C}}}{b_{\mathrm{Ext}}^{*}} < 1 - \max_{u \in \mathcal{C}} |\mathbf{r}_u|$$

then all remaining out-going interactions can be skipped. It is important that all internal interactions are evaluated before the criterion is applied. Otherwise it may incorrectly prevent some short-cut to be created in the component, if out-going interactions would also contribute to building the short-cut.

6.4 Parallelization

The bulk of the work in Algorithm 1 is done in the loop over all pairs of infrastructure nodes which constructs the request vectors. The loop can be parallelized since no iteration of it affects any other. It is not until the last loop any changes are made to the system.

7 Discussion

In the previous sections details of the model are studied. In the following sections a step back is taken when looking at the model to see how the model relates to different aspects of infrastructure.

7.1 Accumulation of Requests

Notably, the model fails to capture a compromise solution in the triangle isosceles system (Figure 5(a)) when the dots represent clusters. The "compromise" $\mathbf{r}_u$ is also a bit unrealistic[5] in the sense that opposite requests cancel each other. An alternative would be to have r_{uv} where v is a neighbor of u and let a request at u increase r_{uv} for several v, but by different amounts depending on how much the deviation from the goal is estimated to cost. However, caution must then be taken to avoid that the same request pushes more than one r_{uv} over the threshold. This idea can be extended further by letting every interaction place requests in an arbitrary number of nodes. While this might make the model handle a compromise solution in the isosceles triangle clusters system it would introduce a whole new level of complexity in the model.

7.2 Calibration, Self-Similarity, b_{ij}, and D_{ij}

The interaction thresholds are scale invariant. Notably the special case of one-cell components can be used to calibrate the model. For clusters and dense active areas, if the aggregation of interactions in b_{ij} balance out the decrease in real interactions due to D_{ij}, then the whole system might exhibit scale invariance like real road networks. Since road networks seem to exhibit self-similarity across a wide range of scales this may say something about how D_{ij} should look. First of course the expression for b_{ij} needs to be found. The mechanism of aggregation in b_{ij} might be empirically hinted by looking at community structures in real urban systems. If D_{ij} can be empirically found instead, then maybe that can be used to hint the form of b_{ij}.

Since b_{ij} represents a community interaction it may be necessary to extend the model with "community centers". A community center would be a cell at the "center of activity" of all the activity added up in b_{ij}. An interaction should then first go to the community center, then to the community center for the goal node, and then finally the goal node. This, however, does not come without its own set of complications.

7.3 Interpretation and Alternative Form of $b_{ij}^{(\text{real})}$

The form of $b_{ij}^{(\text{real})}$ is not obvious. Due to memory limitations x_{ij} is an unmanageable quantity. That draws the attention towards (2). If both the multiplicative and additive part of (2) is included in $b_{ij}^{(\text{real})}$ then the interactions in a two-cell

[5] Unrealistic in a relative sense. Representing individual requests as vectors is unrealistic to begin with; even having individual cells making requests at different locations is unrealistic. At best one can hope that the model as a whole represents a phenomenon that is realistic.

system, assuming $b_{ij}^{(\mathrm{real})} = b_{ij}$, would be

$$b_{ii} = x_i \left(q \frac{x_i}{x_i + a_{ij} D_{ij} x_j} + (1-q) \frac{1}{1 + a_{ij} D_{ij}} \right)$$
$$b_{ij} = x_i \left(q \frac{a_{ij} D_{ij} x_j}{x_i + a_{ij} D_{ij} x_j} + (1-q) \frac{a_{ij} D_{ij}}{1 + a_{ij} D_{ij}} \right)$$

and similarly for b_{ji} and b_{jj}, with the limits

$$\lim_{x_i \to \infty} \begin{pmatrix} b_{ii} & b_{ij} \\ b_{ji} & b_{jj} \end{pmatrix} = \begin{pmatrix} \infty & \infty \\ x_j \left(q + (1-q) \frac{a_{ij} D_{ij}}{1 + a_{ij} D_{ij}} \right) & x_j (1-q) \frac{1}{1 + a_{ij} D_{ij}} \end{pmatrix}.$$

The interpretation of this b_{ij} is somewhat difficult. For instance, there is infinite interaction to j, which itself has finite activity.

In comparision (9) has a more appealing interpretation. If b_{ij} is seen as the need to interact, then the dominant cell has almost no need to interact with the other cell. The dominant cell holds so much activity that it is self-sufficient. The other cell however completely depends on the dominant cell. The analogy with a major city and its suburbs seems clear, even though the real world is not as extreme.

On the other hand, a_{ij} could be left out altogether in $b_{ij}^{(\mathrm{real})}$. The measure would then rather represent how trade and traffic would look if all cells were connected by the infrastructure. Instead of letting the infrastructure follow the current interactions it would in some way try to match a desire of increased interactions.

7.4 Accounting for Congestion When Selecting Trade Nodes

A simple one-dimensional example will be used to illustrate one of the effects of adding a capacity property to the planar road network. Let the nodes be mapped to a one-dimensional lattice such that all nodes, except two, have two neighbors and a road segment between all neighbors. Since the lattice is one-dimensional the path between two nodes is trivial. Let the travel time through a node k be

$$t_k = \max \left\{ 1, \left(\frac{f_k}{C} \right)^{\beta} \right\}$$

where C and β are properties of the road and

$$f_k = f_{kk} + \sum_{i<k} \sum_{j \geq k} (f_{ij} + f_{ji}). \tag{11}$$

f_{ij} is the flow going from node i to node j and f_k is the total amount of traffic in node k. The total travel time between nodes i and j, assuming $j \geq i$, is

$$t_{ij} = \sum_{k=i}^{j} t_k.$$

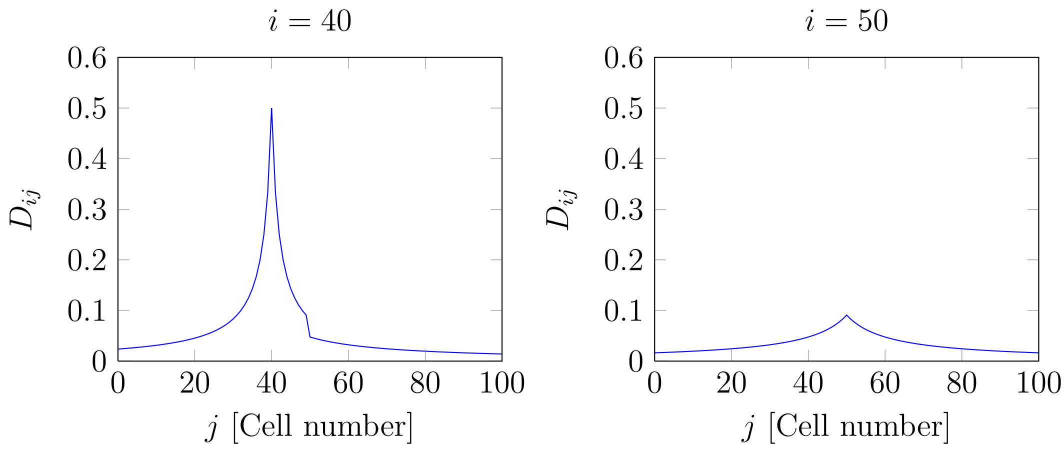


Figure 8: Plots showing the effect of having different travel times through different cells in a one-dimensional model. $D_{ij} = (1 + t_{ij})^{-1}$ and t_{ij} is the travel time between i and j. The travel times through each of the 101 nodes are unity except in node 50 where it is ten times as large. The left panel shows D_{ij} from the perspective of node 40 and a clear drop is seen at node 50. The right panel is from the perspective of node 50 and illustrates the flattening effect.

The travel time can be used instead of the Euclidean distance in the interaction strength (3),

$$D_{ij} = (1 + ct_{ij})^{-\alpha} .$$

The interaction strength is in turn used in the secondary selection (2) and in the interactions (4).

Assuming that more active nodes generate more flow, then the travel time from a very active node to all nodes, including the node itself, will be large and t_{ij} will be relatively flat due to the large "start" cost. In particular, it means that D_{ij} will have a small value for all j, making it too relatively flat as a function of j,

$$D_{ij} = (1 + c(t_i + t_{i+1,j}))^{-\alpha} , \qquad j > i.$$

It is the shape, not the absolute values, that are relevant in D_{ij} due to the normalization in (2) and (4).

Thus, the effect is that spatiality loses value. The additional cost of traveling far is small compared to the minimum travel cost so basically any node can be selected as the destination. An example of this is illustrated in Figure 8.

In two dimensions the effect is not as immediate, but qualitatively the effects above should also be valid in two dimensions. An active cell will sooner or later have enough traffic to/from it that all roads out of the cell will be congested.

There is not much to do with this effect, since in a planar model one cannot circumvent the congestion once all roads are congested. However, if the network is not restricted to being planar then congestion could instead be a driving force to create new faster ways of leaving a cell and ending up farther away, *e.g.* subways for fast inter-city transportation or air travel between cities.

If the internal traffic is not included in the flow (11) then the effect is reversed. Selecting the start node as the destination is then much cheaper than selecting any other node. This simply serves as an isolation of trade and effectively minimizes the mechanism of secondary selection in the trade model.

Calculating traffic flows is a computationally expensive operation. Since it does not seem to bring a qualitative dimension to the model, it seems reasonable to exclude it from the model.

7.5 Planning

There can be some peculiar results such as those in Figure 7. An idea could be to implement some form of planning. The easiest way to do a first-order planning is probably to simply run the system forward for a couple of iterations and keep track of where the last road segment of the first continuous road piece was built, for each interactions. Then the system is restored, and the interactions now seek to get to where the last piece of the first road was built. However, there are two major problems with such a planning. The first is that the run time is increased proportionally to the number of iterations that the planning does. The second is that the new target nodes are not practically storable in the working memory since it is a quantity relating every node to every other node. A more clever form of planning is thus needed.

7.6 Open-Endedness

The model is open-ended with regards to underlying geography. The $\underset{\sim}{w}$ and $\underset{\sim}{w}'$ matrices need not be set up using unit costs. Each edge can have its own associated costs. An extreme case would be water for which the transportation costs would be infinite and construction costs finite but large. Care however needs to be taken to the interpretation of $b_{ij}^{(\mathrm{real})}$ and the model assumption that distance approximates travel time.

8 Conclusions and Future Work

The major work in the thesis is a road growth model. The road growth model uses a cost minimization approximation to decide where new roads are needed. The decisions to grow more road are locally evaluated and there is no global planning. Co-operation naturally occurs in the model when two disconnected road networks try to connect, but a level of explicit local co-operation is needed to avoid two artefacts of the road growth algorithm; to achieve co-coperation to grow short-cuts and to make the model qualitatively insensitive to spatial resolution. Notably the linear cost functions together with the local aggregation of interactions make for a scale invariant system. Exploring the aggregation might give a clue to why road networks exhibit self-similarity.

A suitable next step is to make an efficient implementation of the model and run it on lattices large enough to study statistical properties such as how the average number of reachable intersections increase with travel distance, area distributions of land enclosed by roads, travel profiles, and other properties mentioned under Section 3.3.

The model is very simplified, and none of the aspects mentioned in Section 3.2 are taken into account. The effects of these simplifications are hard to predict without an implementation of the model that produces quantitatively useful data.

The model in itself does not lend itself to making predictions or evaluating policies. At best one can hope that the model can be used to gain some knowledge about land use and road growth in general.

References

C. Andersson, A. Hellervik, K. Lindgren, A. Hagson, and J. Tornberg. Urban economy as a scale-free network. *Phys. Rev. E*, 68(3):036124, Sep 2003.

C. Andersson, K. Frenken, and A. Hellervik. A complex network approach to urban growth. Papers in Evolutionary Economic Geography (PEEG) 0505, Utrecht University, Section of Economic Geography, Feb. 2005a.

C. Andersson, A. Hellervik, and K. Lindgren. A spatial network explanation for a hierarchy of urban power laws. *Physica A: Statistical Mechanics and its Applications*, 345(1-2):227 – 244, 2005b. ISSN 0378-4371.

A. L. Barabasi. Mean field theory for scale-free random networks. *Physica A Statistical Mechanics and its Applications*, 272:173–187, Oct. 1999.

M. Barthélemy and A. Flammini. Co-evolution of density and topology in a simple model of city formation. *Networks and Spatial Economics*, 9:401–425, 2009. ISSN 1566-113X. 10.1007/s11067-008-9068-5.

D. Braess, A. Nagurney, and T. Wakolbinger. On a paradox of traffic planning. *Transportation Science*, 39(4):446–450, nov 2005.

G. Csányi and B. Szendrői. Fractal–small-world dichotomy in real-world networks. *Phys. Rev. E*, 70(1):016122, Jul 2004.

R. C. Dalton. The secret is to follow your nose - route path selection and angularity. *Environment and Behavior*, 35(1):107–131, Jan. 2003.

B. Jiang and C. Claramunt. Topological analysis of urban street networks. *Environment and Planning B: Planning and Design*, 31(1):151–162, January 2004.

V. Kalapala, V. Sanwalani, A. Clauset, and C. Moore. Scale invariance in road networks. 73(2), Feb. 2006.

S. Lämmer, B. Gehlsen, and D. Helbing. Scaling laws in the spatial structure of urban road networks. *Physica A Statistical Mechanics and its Applications*, 363:89–95, Apr. 2006.

L. J. LeBlanc, E. K. Morlok, and W. P. Pierskalla. An efficient approach to solving the road network equilibrium traffic assignment problem. *Transportation Research*, 9(5):309 – 318, 1975. ISSN 0041-1647.

I. Milchtaich. Network topology and the efficiency of equilibrium. *Games and Economic Behavior*, 57(2):321 – 346, 2006. ISSN 0899-8256.

S. Porta, P. Crucitti, and V. Latora. The Network Analysis of Urban Streets: A Dual Approach. *eprint arXiv:cond-mat/0411241*, Nov. 2004.

A. Turner. From axial to road-centre lines: a new representation for space syntax and a new model of route choice for transport network analysis. *Environment and Planning B: Planning and Design*, 34(3):539–555, May 2007.

I. S. R. M. R. E. United States Department of Agriculture, Forest Service. Cost estimating guide for road construction, June 2009.

G. Valiant and T. Roughgarden. Braess's paradox in large random graphs. In *Proceedings of the 7th ACM conference on Electronic commerce*, EC '06, pages 296–305, New York, NY, USA, 2006. ACM. ISBN 1-59593-236-4.

H. Youn, M. T. Gastner, and H. Jeong. Price of Anarchy in Transportation Networks: Efficiency and Optimality Control. *Physical Review Letters*, 101(12): 128701–+, Sept. 2008.