

Commercial Evolution Simulation

Carlos Soriano[†] & Gustavo Patow[‡]

ViRVIG-UdG

Abstract

Simulating the evolution of urban landscapes is a challenging objective with a large impact not only for Computer Graphics (for its applications in the filming and gaming industries), but also for urban planning, economical and historical studies, urban physics, and many other. However, this target has remained elusive because of the large complexity implied by urban structures and their evolutions. We present a system that aims at simulating the evolution of the commercial structure in a modern city. In particular, given an initial distribution of shops, it studies the evolution when larger commercial areas, like malls, are introduced. This is computed using the Huff model as a measure of the attraction each commerce has on potential consumers, and an agent-based simulation to determine how these aspects affect their choices. Then, after a given simulation time, the system decides whether the shop has retained an income such that it can continue operating, or has gone bankrupt. Our system is used to study the evolution of the commercial structure of Barcelona city over the last century.

1. Introduction

Procedural urban modeling has presented us with astonishing results over the last decade, starting with the seminal work by Parish and Muller [PM01] and Muller et al. [MWH*06], and continuing with the recent advances in acquisition [MWA*12], non-regular modeling [LCOZ*11], user interfaces [Pat12], among others.

However, in spite of all those improvements, several problems remain open [PBP14], one of the most important ones is simulating the evolution of urban landscapes over time. With only a few exceptions [WMWG09, BWK14], this topic has barely been touched, in spite of its crucial importance for history and archeology, urban planning, socio-economical studies, and many other social-related disciplines.

Among these unexplored aspects, the problem of simulating the evolution of the commerce structure in a city is a prominent one, as it is attractive for being computationally tractable and crucial for socio-economic studies. But this study has applications that are broader than a pure social analysis, as the resulting distributions can be used to also model its appearance over time, which is interesting for computer graphics because of its applications to film and videogames, two of the leading industries in the field.

[†] Currently at Red Hat. carlos.soriano89@gmail.com

[‡] dagush@imae.udg.edu

2. Previous Work

After the seminal works by Parish and Muller [PM01] and Muller et al. [MWH*06], based on the idea of L-systems, several highly relevant research papers related to city modeling appeared. An approach for street network generation based on templates was proposed by Sun [SYBG02]. Chen et al. [CEW*08] proposed using tensor fields to guide the generation of street graphs. Later, Ridorsa and Patow [RP10] presented an interactive flow-oriented metaphor that is both simpler to use and to understand. For more information, we refer the reader the surveys from Kelly and McCabe [KM06], Watson et al. [WMV*08], Vanegas et al. [VAW*10] or Smellik et al. [STBB14].

For the specific case of the temporal evolution of cities, we can mention the work by Weber et al. [WMWG09], who grew streets based on an L-system-like approach, but used traffic simulations to validate the streets to add. This work was later extended by Benes et al. [BWK14], who took into account a city's neighborhood for the traffic simulations. Vanegas et al. [VGDA*12] used inverse modeling techniques to simulate and predict the general behavior of an urban environment.

As opposed to geometric urban modeling, which is purely Computer Graphics-oriented, behavioral urban modeling is intended for decision-making regarding urban policies in current and future urban areas. Two of the most important examples are the works of Alkheder et al. [AWS08] and

Wadell [Wad02]. Vanegas et al. [VABW09] approach consists in closing the loop between behavioral and geometric modeling. This is accomplished by subdividing the urban space into a regular grid of cells, each one of which is assigned a set of variables that control the distributions of population and jobs, land values, road network, parcel shape and building geometry. These variables can be globally or locally increased, decreased or constrained. Their behavioral framework – which is based on a simplification of Urban Sim [Wad02] – and their geometrical modeling engine produce a single dynamical system that seeks behavioral and geometric equilibrium after each user-specified change. That system, however, lacks of some elements of behavioral modeling that could be provided by a traffic simulation.

Traffic simulations can be classified in several ways, the most basic one being the division between continuous- and discrete-time approaches, and between microscopic (describe the system entities and their interactions at a more detailed level of detail), macroscopic (describe entities and their activities and interactions at a coarse level of detail), and mesoscopic simulations (a midpoint that usually represents most entities at a high level of detail but describe their activities and interactions at a much lower level). One of the oldest and most well known cases of the use of simulation in theoretical research is that of Gerlough and Huber [GH75], in which a differential equation governs the movement of each vehicle. Recently, the concept of continuum macroscopic traffic simulations and treating cars as gas particles in the context of computer graphics has been explored by Sewall et al. [SWML10]. Their simulation describes the flow of traffic through a system of nonlinear hyperbolic conservation laws that represent traffic as a continuum along lanes, following the work by Lebaque [LMHS07] and augmenting the model proposed by Aw and Rascle [AR00] and Zhang [Zha], extended to be able to handle lane changes and merges, as well as traffic behaviors due to changes in speed limit. However, their approach is heavily limited in the scope of urban planning, given that their vehicles “flow” without any intentionality, and is also limited in the type of roads involved, which can only be of highway-class roads without intersections. Two of the most important microscopic simulation systems available are MATSim [MBH*10] and SUMO [KBW06], both of them vehicle-based. Both systems are able to simulate large road networks using agent-based microsimulations. MATSim’s pipeline, starts with the generation of daily plans for each agent. These plans are based on information from several sources like census data and network data, and consist of a series of desired activities and paths (or “legs”) between them. To calculate the optimal paths to carry these plans out, MATSim offers modules based on known algorithms like Dijkstra and A*, as well as an optimization of the latter named Landmarks-A*. Then, a simulation based on genetic algorithms is performed, which iterates over three steps: execution of the plans, scoring them and replanning according these scores. Finally, an analysis

stage is performed. Their toolbox provides several modules that can be combined in different ways to perform each one of these steps. It can take up to 3.2 days to simulate the traffic for a day in all of Switzerland. Our approach follows an analog path, by simulating the behavior agent by agent, much in the same way as MATSim or SUMO work.

Another quite related field is mathematical modelling and the use of models in planning in relation to all aspects of cities and regions - including demography, economic input-output modelling, transport and locational structures. In particular, the applications of dynamical systems theory in relation to modeling the evolution of urban structure which are the foundations of a comprehensive theory of urban dynamics described in Complex spatial systems, as described, for instance, by Dearden and Wilson [DW11] [DWA15].

3. Commerce Simulation

In this section we will introduce our commerce simulation model and all the details of its implementation.

3.1. Commercial Model

The distribution of shops and other commercial entities changes over the time due to different factors, like competition, population distribution changes or even consumer preferences. In Urbanism, different hypothesis have been formulated to study those interactions, used to statistically predict which commercial entity is more likely to be chosen.

In this paper we used the model developed by Huff [Huf63], which is a spatial interaction model that calculates gravity-based probabilities of consumers at each origin location patronizing each store in the store dataset. From these probabilities, sales potential can be calculated for each origin location based on disposable income, population, revenues, or other variables. In our case, this importance S_j was set as the taxes paid by the store owner, which we assume proportional to its overall activity. The probability values at each origin location can optionally be used to generate probability surfaces and market areas for each store in the study area. The Huff formula is:

$$P_{ij} = \frac{S_j^\alpha T_{ij}^\beta}{\sum_{k=1}^n S_k^\alpha T_{ik}^\beta} \quad (1)$$

where the different symbols are defined in table 3.1.

To compute the distance between each agent and any given commercial entity we used the well known *urban distance*, although any other measure (e.g., Euclidean) can be used. Basically, urban distance is the distance between two points taking into account the shortest path following the streets of the city, which we implemented using the well known A* algorithm. This distance is computed differently whether we are considering pedestrians, who can traverse any street in any of its two directions; or cars, which are

P_{ij}	Probability to go from entity i to j
S_j	Importance of a commercial entity
T_{ij}	Travel time or distance from commercial entity i to commercial entity j
α	Sensibility to the commerce importance (empirical value, 1)
β	Sensibility to time travel or distance (empirical value, -2)
n	Total number of commercial entities

Table 1: Symbols used for the Huff model. The values for α and β are the ones recommended by Huff [Huf63] (1 and -2 respectively).

bound to follow street and lane directions. Although the implementation in both cases does not change much, the consequences as to where a consumer will be more leaning towards could drastically change in some complex street networks. In the case of pedestrians, we ignored the street directionality, while for cars they were taken into account. It is important to note that, in the later case, T_{ij} could be non-symmetric, as the distance to go from entity i to entity j could be different as we travel the other way round because of the different street directionality.

3.2. Simulation

The first step in our simulation is to compute, at any given moment in time, the economic performance of a set of commercial entities. For that, we generate a (large) number of agents, and we let them decide the entity they will go to acquire goods with a random selection proportional to the Huff-calculated probability value. Then, we simply accumulate, for each entity, the number of agents that selected it to get supplies. As a result we have, for each entity, a floating point value that represents its income (or a factor proportional to this value).

The second stage of our algorithm then is executed for a different time, with a new set of commercial entities that include subset of the ones presented at the previous time (some shops might have been closed for any possible reasons), some newly created ones. These new entities will have their importance evaluated and stored along with the ones previously existing. If, after a simulation step is computed, the income of the entity gets lower than a given user-defined threshold (in our case, defined as a percentage of the income in the previous iteration), the entity goes bankrupt. For newly created entities, this value is taken as a fixed constant proportional to its initial importance. The decision of creating a new entity is taken, in our implementation, solely based on the evolution of the population density of in a given area. If the population increased, new shops are sampled proportional to the change in population local to that area, and the importance is assigned with a gaussian distribution using the average and

variance of the existing shops in the area. Closing shops is accounted for in a natural way, as already described.

3.3. Data acquisition

This process was delicate and quite error-prone, as we manually acquired the information for each shop at the specified years from the local city council archive, which is not in digital form. Actually, the situation is even worse as the information to gather is hand-written in old official registry books. As a result, this manual step resulted in a list of shops and associated information for each of the studied years. However, as a result, it had several inaccuracies and other problems that came from this manual setup. In this section we are going to explain some of the algorithmic steps taken towards cleaning our input information.

For the real data of commercial entities we used a CSV file with a few necessary fields like the street name where the entity is, the importance of the entity, and an optional field for a location. To know the location (latitude and longitude coordinates) of each entity, we used Google Maps. The process is quite simple: We query Google Maps for the coordinates of a given street name, and if there is a result good enough, Google Maps returns the latitude-longitude pair.

However, the process gets much more involved as our data comes with the logical typos and wrong street names resulting from a manual acquisition based on hand-made annotations. Usually, we can safely assume typos are introduced by the people who put the data in the CSV file. However, wrong street names are more related to the street name changes over the time and, in this particular case, to a specific event in Barcelona city: The original street names were in Spanish in the past due to law at the moment of the recording. However, with the arrival of the democracy, streets are now named using the local language, Catalan, and so are stored in Google Maps. Thus, any query of an old Spanish name would result in either an empty answer or would return a place that might not even belong to Spain. Filtering by the city bounding box in latitude-longitude may prevent the system from adding these wrong coordinates, but the first kind of problems had to be solved with a specific approach. To solve this, we used Google Translate to ensure all street names are in the same language (i.e., Catalan). This solved most of our language problems, providing results that Google Maps could handle correctly. As a fall-back, if neither of the methods listed above works, we allow the user to manually add the coordinates at the CSV file, so important markets are not lost because of the source data mismatches.

4. Results

In this section we are going to present our results for this ongoing project. We firstly evaluate the performance of our hypothesis by testing the Huff model predictive power, and we studied what we called the natural evolution of the

commerce in a city, following the population distribution changes and the occasional apparition or disappearance of some shops. Then, we proceeded with the simulation of the real events that happened at the beginning of the century, when several national markets were created by the new regulations in Spain. This resulted in many smaller shops to go bankrupt, so we decided to test our model in this scenario.

4.1. Testing Huff model

Originally, the Huff model was designed to provide a numerical measure of the likelihood of a potential customer to go to one shop or another among a given number of options. It was never intended as a performance measurement tool, and much less as an evolution-predictive analysis tool. So, to verify its suitability in our algorithms, we started with a set of shops uniformly distributed on a homogeneous, procedurally generated city, and studied the number of customers that went to each shop, assuming also a uniform distribution in the population. This resulted on a reasonably well-shaped distribution of customers for each shop, with an average being roughly equal to the number of inhabitants samples divided by the number of shops to select.

As a second step, we introduced one large entity, with a high importance, and repeated the simulation to see the difference that this big entity introduced in the model. If the simulation parameters are correct, we would observe a drop of the agents for each shop following a linear relationship. In Figure 1 we can see a plot of the result given the distance between a small entity and the big one.

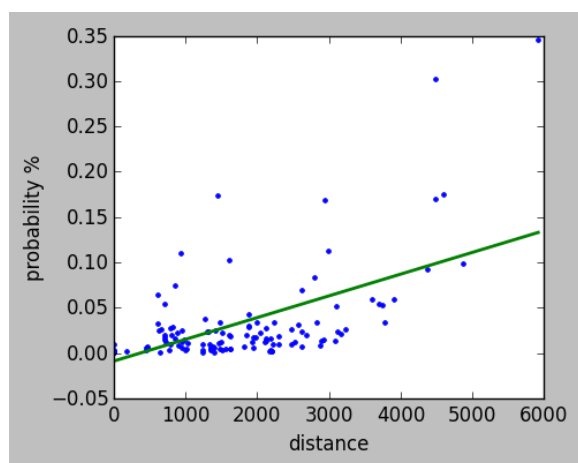


Figure 1: Testing Huff model on a controlled scenario: relationship between the agent shop selections and the distance to a national market (not included).

For the simulation we used real data of the distribution of shops and similar commercial entities in the city of Barcelona in 1897, where the presence of large commercial

entities was minimal and small shops were the general kind of entities. We generated a distribution of customers (agents) proportional to the population density. For that, we embedded the city in a grid, and sampled agents proportionally to the density for the respective year in each cell. We then selected the street points closer to the sampled positions, and projected the agents to these street positions, rejecting the invalid ones (e.g., out of the map/street or unconnected streets) and re-assigned the agents to a different position. We applied the Huff model for each agent and for each commercial entity, calculating its probabilities to go to a specific entity. Using a cumulative distribution function (CDF), we randomly selected which entity will be the agent destination. With this information, we accumulated the total number of agents for each destination, obtaining the base income for each commercial entity. Now, for any other year we would like to simulate, we can compute the difference of the income for that entity for the given year and the base income, and decide whether this shop continues working or, if the income fails below a user-defined threshold (in our case, 60% of the base income), it becomes bankrupt. In Figure 2 we can see the situation in Barcelona at year 1897.



Figure 2: Barcelona commercial entities at 1897. Entity scales are on par with their importances.

Once this initial information was setup, we added the large national markets that were created at Barcelona in 1932, when a law enforced to build that kind of markets, which normally were several times (between 4 and 50 times) larger than a normal commercial entity. This caused small shops to go bankrupt if they were close to the new market, or if they were not large enough to avoid the agent loss by the large market. In Figure 3 we can see the location of the new markets created in 1932 in relation to the original ones.

After the simulation, the resulting distribution can be seen in Figure 4, where many small shops went bankrupt (black dots) because of the decrease in income as a result of the apparition of the large markets. As can be seen, the introduction of these markets produces, according to the simulation,



Figure 3: Barcelona national markets, 1932. Pink dots correspond to regular entities and green dots to national markets. Dot size is proportional to its importance.

the elimination of almost half of the original shops. This can be understood more clearly if we study the plot of the attractiveness of a uniform distribution of shops when a single national market is introduced, as shown in Figure 5. As can be seen, the introduction of each national market results in a depression of the attractiveness in close areas, with a recovery at farther distances where the influence of the national market diminishes (remember that we use Huff's quadratic scheme). This results in an overall probability distribution that can be depicted in the density plots at Figure 6, for both before and after this introduction.



Figure 4: Barcelona commercial entities at 1932 with bankrupt represented as black dots

If we compare these theoretical results, we would see that the model predictions are statistically right (i.e., on average) at a distance far from the markets. However, when comparing the situation with what really happened until 1932, we would see an important *increase* of the number of shops in

the immediate vicinity of the markets. The explanation is that these new shops *benefited* from the immediacy to the market, as people would not perceive it as a problem to go out of the market and walk less than a block for supplies. The Huff model is unable to predict this "stickiness" effect, which should be left as future work for our model.

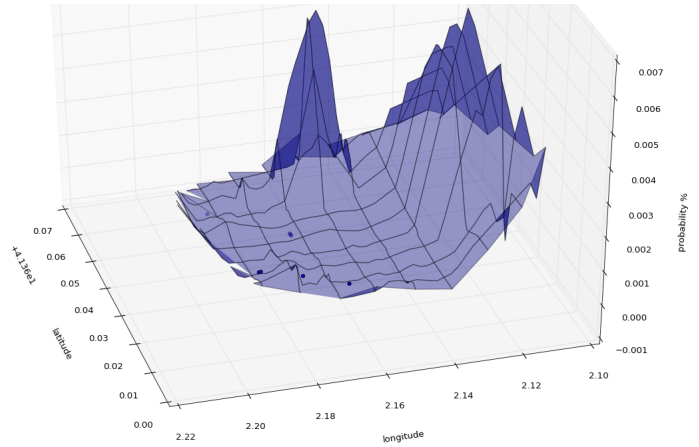


Figure 5: A plot showing the spatial probability of agents of going to a regular shop.

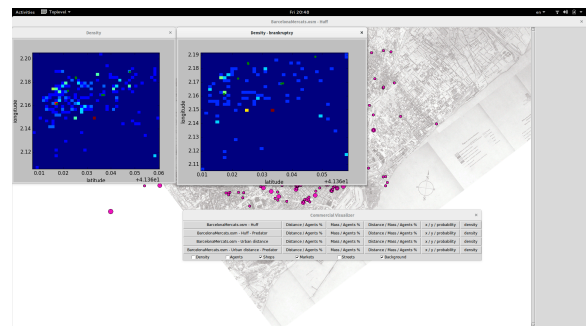


Figure 6: A plot showing the density of regular shops before and after the introduction of the national markets.

5. Conclusion

We have presented an algorithm to predict the evolution of shops and markets in a modern city over time. The model is based on the Huff gravitation estimation, and a simulation given in terms of the urban distance from prospective customers to all the shops. The algorithm then used sampling techniques to generate a population of agents/customers that selected a potential commercial entity and contributed to its final income result. This simulation allowed us to partially predict the evolution of the commercial market of the city of Barcelona from the situation in 1897 to the one in 1932, taking into account the introduction of the national markets.

Our system correctly predicts the average behavior at large distances from the markets, but fails to predict the "stickiness" effect resulting on the attractiveness of the markets over themselves, but also over the surrounding area. It is left as future work to add this behavior in our models.

One line for future improvement is to deepen the computations of shop quality, opening, and closing circumstances, to reflect more real-life situations. Another future avenue of research is to explore the connection with dynamical systems theory in relation to modeling the evolution of urban structure.

Acknowledgements

We would like to thank the anonymous reviewers for their valuable comments, as their insights really helped improved this paper. We also would like to thank Nadia Fava for pointing us to the Huff model and for providing some data to validate the model. This work was partially funded by the TIN2014-52211-C2-2-R project from Ministerio de Economía y Competitividad, Spain.

References

- [AR00] AW A., RASCLE M.: Resurrection of "second order" models of traffic flow. *SIAM Journal of Applied Mathematics*, 60 (2000), 916–938. [2](#)
- [AWS08] ALKHEDER S., WANG J., SHAN J.: Fuzzy inference guided cellular automata urban-growth modeling using multi-temporal satellite images. *Int. J. of GIS* 22, 11-12 (2008). [1](#)
- [BWK14] BENEŠ J., WILKIE A., KŘIVÁNEK J.: Procedural modelling of urban road networks. *Computer Graphics Forum* 33, 6 (2014), 132–142. [1](#)
- [CEW*08] CHEN G., ESCH G., WONKA P., MÜLLER P., ZHANG E.: Interactive procedural street modeling. *ACM Trans. Graph.* 27, 3 (2008). [1](#)
- [DW11] DEARDEN J., WILSON A.: A framework for exploring urban retail discontinuities. *Geographical Analysis* 43, 2 (2011), 172–187. [2](#)
- [DWA15] DEARDEN J., WILSON A., ALAN WILSON J.: *Explorations in Urban and Regional Dynamics*. Regions and Cities. Taylor & Francis, 2015. [2](#)
- [GH75] GERLOUGH D., HUBER M.: Traffic flow theory. a monograph. *TRB Special Report 165* (1975). [2](#)
- [Huf63] HUFF D.: A probabilistic analysis of shopping center trade areas. *Land Economics* 39, 1 (1963), 81–90. [2, 3](#)
- [KBW06] KRAJZEWICZ D., BONERT M., WAGNER P.: The open source traffic simulation package sumo. In *Robocup* (2006). [2](#)
- [KM06] KELLY G., MCCABE H.: A survey of procedural techniques for city generation. *ITB Journal* 14 (2006), 87–130. [1](#)
- [LCOZ*11] LIN J., COHEN-OR D., ZHANG H., LIANG C., SHARF A., DEUSSEN O., CHEN B.: Structure-preserving re-targeting of irregular 3d architecture. *ACM Trans. Graph.* 30, 6 (Dec. 2011), 183:1–183:10. [1](#)
- [LMHS07] LEBAGUE J., MAMMAR S., HAJ-SALEM H.: The aw-rasclle and zhang's model: Vacuum problems, existence and regularity of the solutions of the riemann problem. *Transportation Research B*, 41 (2007), 710–721. [2](#)
- [MBH*10] MEISTER K., BALMER F., HORNI M., REISER M., WARAICH R., AXHAUSEN K.: Large-scale agent-based travel demand optimization applied to switzerland, including mode choice. In *12th World Conference on Transportation Research* (2010). [2](#)
- [MWA*12] MUSIALSKI P., WONKA P., ALIAGA D. G., WIMMER M., VAN GOOL L., PURGATHOFER W.: A survey of urban reconstruction. In *EUROGRAPHICS 2012 State of the Art Reports* (May 2012), EG STARS, Eurographics Association, pp. 1–28. [1](#)
- [MWH*06] MÜLLER P., WONKA P., HAEGLER S., ULMER A., VAN GOOL L.: Procedural modeling of buildings. *ACM Trans. Graph.* 25, 3 (2006), 614–623. [1](#)
- [Pat12] PATOW G.: User-friendly graph editing for procedural modeling of buildings. *IEEE Computer Graphics and Applications* 32 (2012), 66–75. [1](#)
- [PBP14] PUEYO X., BOSCH C., PATOW G. A.: On improving urban environment representations. *Frontiers in Robotics and AI*, 17 (2014). [1](#)
- [PM01] PARISH Y. I. H., MÜLLER P.: Procedural modeling of cities. In *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques* (2001), pp. 301–308. [1](#)
- [RP10] RIDORSA R., PATOW G.: The skylineengine system. In *XX Congreso Español De Informática Gráfica, CEIG2010* (2010), pp. 207–216. [1](#)
- [STBB14] SMELIK R. M., TUTENEL T., BIDARRA R., BENES B.: A survey on procedural modelling for virtual worlds. *Computer Graphics Forum* 33, 6 (2014), 31–50. [1](#)
- [SWML10] SEWALL J., WILKIE D., MERRELL P., LIN M.: Continuum traffic simulation. *Proceedings of EUROGRAPHICS* (2010). [2](#)
- [SYBG02] SUN J., YU X., BACIU G., GREEN M.: Template-based generation of road networks for virtual city modeling. In *VRST '02: Proceedings of the ACM symposium on Virtual reality software and technology* (New York, NY, USA, 2002), ACM, pp. 33–40. [1](#)
- [VABW09] VANEGAS C. A., ALIAGA D. G., BENEŠ B., WADDELL P. A.: Interactive design of urban spaces using geometrical and behavioral modeling. *ACM Trans. Graph.* 28, 5 (Dec. 2009), 111:1–111:10. [2](#)
- [VAW*10] VANEGAS C. A., ALIAGA D. G., WONKA P., MÜLLER P., WADDELL P., WATSON B.: Modelling the appearance and behaviour of urban spaces. *Comput. Graph. Forum* 29, 1 (2010), 25–42. [1](#)
- [VGDA*12] VANEGAS C. A., GARCIA-DORADO I., ALIAGA D. G., BENES B., WADDELL P.: Inverse design of urban procedural models. *ACM Trans. Graph.* 31, 6 (Nov. 2012), 168:1–168:11. [1](#)
- [Wad02] WADELL P.: Urbansim: Modeling urban development for land use, transportation and environmental planning. *Journal of the American Planning Association* 68, 3 (2002). [2](#)
- [WMV*08] WATSON B., MÜLLER P., VERYOVKA O., FULLER A., WONKA P., SEXTON C.: Procedural urban modeling in practice. *IEEE Computer Graphics and Applications* 28 (2008), 18–26. [1](#)
- [WMWG09] WEBER B., MÜLLER P., WONKA P., GROSS M. H.: Interactive geometric simulation of 4d cities. *Comput. Graph. Forum* 28, 2 (2009), 481–492. [1](#)
- [Zha] ZHANG H.: A non-equilibrium traffic model devoid of gas-like behaviour. *Transportation Research B*, 36, 275–290. [2](#)