

How to collect data to simulate the dynamic of trains-passengers' interaction

Fatma Elleuch (fatma.elleuch@reseau.sncf.fr), Stéphanie Donnet (stephanie.donnet@reseau.sncf.fr)
SNCF Réseau, 174 Avenue de France, 75013 Paris

Axel Buendia(axel.buendia@lecnam.net), Fatma Elleuch (fatma.elleuch.auditeur@lecnam.net), Stéphane Natkin (stephane.natkin@lecnam.net)
CNAM CEDRIC, 292 rue St Martin 75003 Paris, France

Charles Tijus (tijus@lutin-userlab.fr)
LUTIN, Cité des Sciences et de l'Industrie
30 avenue Corentin Cariou - F-75930 Paris Cedex 19

Abstract

This paper presents a motivation-based model in order to explore crowd behavior. The case study is about what motivates the decision processes of passengers about choice of location on the station platform for ingressing and egressing trains. The goal of the research is twofold: to establish a cognitive generic crowd behavior modeling method and to respond to a major challenge of public transportation: to reduce dwell time to ensure a high level of service.

We first introduce motivation-based modeling for the simulation of the dynamics of numerous cognitive agents and report the collection of passengers' dynamics that was done through an extensive survey observation. Most significant variables were then extracted from factor analysis to compose and distinguish six main motivation based strategies that are to be used for the simulation of crowd behavior in the train station.

Discussion is about the advantages of motivation-based simulation in terms of robustness and adaptability and conclusion about how Artificial Intelligence, Cognitive Psychology and Data Science operate together to model such complex systems.

Keywords: motivation-based model, crowd behavior, dwell-time, generalized linear model

Introduction

If the dynamics of crowd behavior are hard to study because people are not reliable witnesses of their own behavior. This is mainly because of implicit motives derived from affective experiences (McClelland et al., 1989) or when constrained by contextual and situational affordances provided by (de Lavalette et al., 2009). Otherwise, people do have cognitively elaborated constructs that sustain their current goal; as motives of their decision-making for acting or behaving in a particular way.

Our research work is about a motivation-based method to analyze and simulate crowd behavior. The case study is the passenger's flows transfer when ingressing or egressing from a train at a given station using a real-time survey: asking here and now people about the motives of their behavior. For the train operator, the goal of the simulation is to find how to modify the station layout and the platform-train interface for minimizing planned dwell time.

We provide below (i) the passengers-train context of the case study used for our applied cognitive science study, (ii) the presentation of the motivation based model for the simulation of multi-agents behavior for crowd analysis, (iii) the method and procedure used for collecting data, modeling behavior on the platform in order to instantiate the simulation model. Main results are six strategies of passengers in station about how to behave while waiting for the train.

Context and description of the case study

In the transportation literature, dwell time is considered as a key variable in performance, reliability and quality of service (Puong, 2000, Hutton, 2013, Fernandez et al., 2015). Our case study relates on the interest on reducing dwell time to increase the train transportation capacity, to avoid congestion situation and, therefore, to improve the level of service.

The case study is the decision-making of passengers about their location on the station platform for ingressing and egressing trains. The train station¹ has an architectural complex station platform characterized by a high density of passengers, especially at the end of the day (the peak hours from 5pm to 8pm). This station has a set of particularities supposed to implement in a multi-agents model. First, it is an island platform. This type of platform is known to generate issues such as overcrowding, especially when two trains are stopping on either side. It also causes multiple crossing passenger flows and overcrowding around stairs and escalators (Yamada et al., 2014). Second, due to traffic regulation, modifications of the served side of the platform happen frequently at the last moment, less than 2 minutes before the train arrival. Passengers are informed through information's board and verbal calls. Thus, in addition of the longitudinal distribution, there is also a transversal distribution of passengers waiting for their train.

¹ Bibliothèque François Mitterrand station

Crowd behavior and motivation-based modeling

Crowd simulation is the process of simulating dynamics of numerous moving entities. Due to the expected high precision of the model, we focus on the agent-based approach as opposed to more macroscopic approaches (Henderson, 1995; Zhang et al., 2009; Helbing et al., 1995), by using a motivation-based model (Spier & McFarland, 1997; Constant et al., 2015).

Agent-based approach serves to model the crowd behavior of autonomous agents in interaction usually by a bottom-up construct that emerges from a set of simple rules that describe agent attributes and behaviors (Macal et al., 2010). This kind of models is in continuous improving and enrichment and is operating within human sciences, providing relevant information on both individual and collective behaviors. Thus, on the one hand, agent-based system captures the multidimensionality of social phenomena (Boero, 2015). On the other hand, motivation-based system allows incorporating a further level in what might be the hierarchical decomposition of the behavior. The mechanisms based on motivational tendencies perform better, are more adaptive, and demonstrate opportunism (Spier & McFarland, 1997).

Within the motivation-based model, each agent shares a set of motives. These motives are computed dynamically and individually according to personal attributes. Those attributes encompass personality traits, preferences, or even initial knowledge and are used to calculate the motivation strength for each agent. In turn, motivations generate the behaviors used for simulation.

To model and simulate our case study that is the behavior of passengers at the platform, information are collected through passengers Real-Time surveys about why currently they are behaving as they do, in order to collect data about people's motives as well as the variable parameters of those motives in the context and situation of the station: entries / exits, conveniences, information boards....

For the simulation of the passengers' behaviors, we use SpirOps Crowd (Buendia, 2003) instantiating its multi-agent based frameworks with passengers' motives in the context of the station with real-time arrivals and departures of trains.

Method

To locate the passengers' position when ingressing or egressing from a train, the platform was divided into 12 zones (Figure 1) delimited by the visible visual landmarks of the platforms pillars that stand at the same distance from each other.

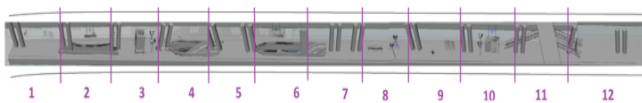


Figure 1: the platform divided into its 12 zones

The survey was conducted the evening from a Monday to Thursday. Participants were 545 passengers recruited in the train station. The inclusion method of participants was to randomly select passengers who were waiting on the platform for their train, having for each zone the same number of respondents (≈ 45). Passengers were asked for their age, motive of displacement, regularity of the trip and their motivations about their longitudinal position (*why this zone rather than the others*) and transversal position (*why this side rather than the other side*).

The mean age of the 545 participants is 39.1 (SD =12.9), with an equivalent men/women proportion (respectively 48.5% and 51.5%). We notice mostly regular users (90.5%) and a predominant motive of displacement that was going back home (94.5%)

Results

Among types of motives, at the first place, we distinguish between intentional (97.4%) and unintentional choices (2.6%). This means that participants motivate their "being there" with arguments.

Thus, we classified these intentional choices distinguishing between three categories (Figure 2). First, is the "Platform" category that encompasses passengers who argue caring about the current situation on the platform (density) and its characteristics (entries position, information board, space, seats). This category gathers different attitudes that could be viewed as a result of the human-environment interactions, whether it is a real-time or a planned decision. Due to its richness, this category will be investigated further in more details.

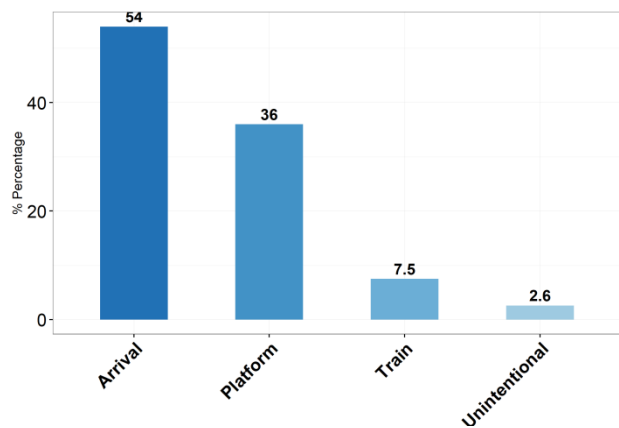


Figure 2: the responses' categories about the longitudinal distribution of passengers (in percent).

Second is the "Train" category that concerns passengers who care about the crowdedness' level inside the train by selecting their favorite vehicle in a comparative way. Mainly, this strategy is to ensure obtaining free seats and guarantees more comfort for the user.

Third is the category "Arrival" that is provided by passengers planning to minimize the distance to the exit at the arrival station.

In some cases, we obtained combinations of these motives. Therefore, we weighted the responses according to the number of combined strategies.

Individual and trip characteristics

For each of these three categories, we computed the mean age (*with standard deviation*), the gender proportion (female / male), the proportion of participants being regular vs. irregular passengers, traveling alone or in a group and the number (N) of concerned passengers in order to analyze how these motives relate to passengers properties.

Table 1: relationships between *Platform*, *Train* and *Arrival* categories of motives with passengers’ properties.

	Platform	Train	Arrival	No Motive
Age	38.2 (<i>sd: 13.3</i>)	43.1 (<i>sd: 11.3</i>)	39.5 (<i>sd: 12.6</i>)	43 (<i>sd: 13.3</i>)
Female	0.47	0.5	0.5	0.21
Regular	0.86	0.94	0.94	0.85
Going home	0.93	0.95	0.98	0.71
Travelling alone	0.95	1.00	1.00	1.00
N	191	40	286	14

The relationships between *Platform*, *Train* and *Arrival* motives and passengers properties were evaluated using the Generalized Linear Model (GLM) (McCulloch & Nelder 1989). Differences in age and travelling regularly were significant: the *train* motives were given mostly by aged people, and the platform motives less given by regular passengers that were also less going home.

We describe below each of the category of strategies according to passengers motives and linked properties.

“Intentional” versus “Unintentional” strategies

Mostly, participants motivate their behavior with arguments. Such an important degree of intentionality could be explained by the special context of the study: as indicated before, experiments took place at the end of the day. Most of the people answering the survey are *Regular* and going back *Home*. Like going to office, going home is a situation that seems to generate also diverse strategies aiming to reduce cost (effort) and increase gain (time, comfort).

“Platform” strategy

Focusing on the different factors and their impact on the intentional decision, we notice that people who adopt “*Platform*” motive are less likely to be regular and slightly more likely to be in groups with other persons compared to the other categories. Distributions of *Regularity* and of traveling *alone* displayed through Figure 3.

There is a slight predominance of the “*Platform*” strategy in all *Irregular* users and *Regular/Group* combination. This result could be inferred to the fact that when a traveler is not

a regular one, s/he didn’t have yet developed strategies about where to be located on the platform. At least, it is more probably a regulatory strategy (regulate position relative to the entrances, step by step) than a planned one (choosing a specific place along the platform). In addition, to be accompanied by another person(s) may divert or decrease the attention about the possible advantages of location because one might be engaged in a social interaction.

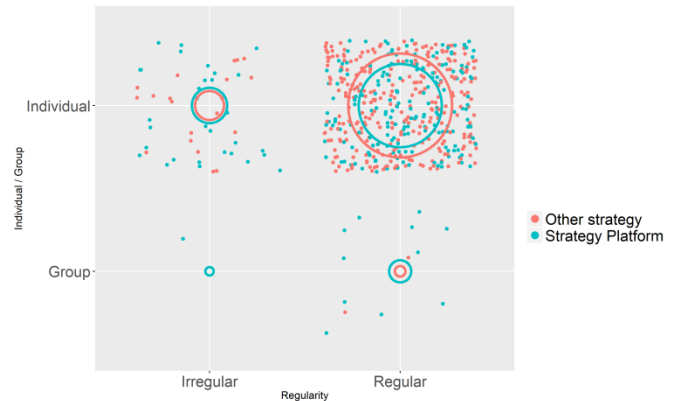


Figure 3: the clustering of passengers according to *Regular* and *Individual* trip with a slight predominance of the “*Platform*” strategy in all *Irregular* users and *Regular/Group* combination.

“Train” strategy:

Passengers that have a “*Train*” motive are more aged than the other categories: age might enhances motivation of searching more comfort (Figure 4).

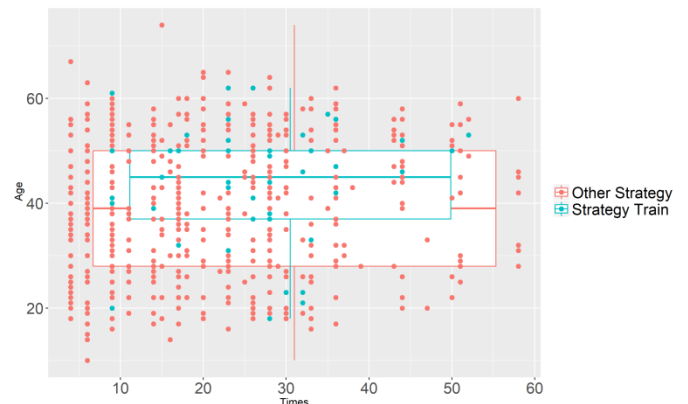


Figure 4: aged participants are more likely to adopt a “*Train*” strategy.

For these statistical analyses, the effect of trip duration wasn’t significant; neither the effect of prior experience of the station (*Regular* vs. *Irregular*), since trying to maximize one’s own comfort aboard a train requires a prior inside train knowledge.

There was a link between age and the number of *train* motives: aged people were more likely to be motivated by comfort inside the train. There was also a link between the number of *train* motives and the expectation to find a free seat inside the train. This may be explained by the following detail: qualitative interaction between surveyor and user reveal user’s expectations about seats release at some stations (e.g., *Savigny sur Orge*, *Time duration = 17mn*). One may also think that the motivation to find a free seat inside the train increases continuously with the trip duration. This is not the case. The number of motives increases up to a threshold (20mn to 40mn) then decreases.

“Arrival” strategy

Finally, passengers of “*Arrival*” category are more likely to be *Regular* and are mostly going home (Figure 5). We emphasize that the *goal* of the trip, which is going home, enhance adopting an “*Arrival*” strategy.

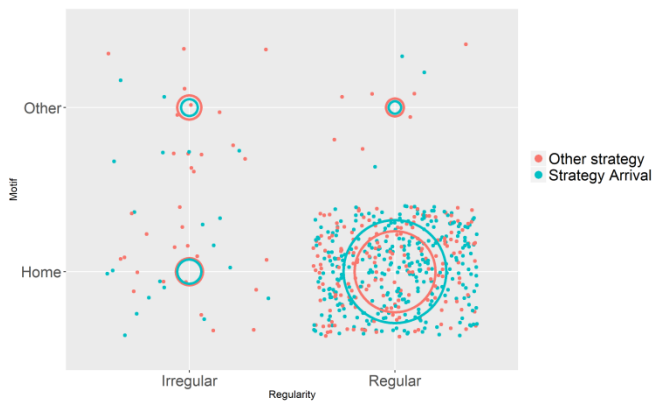


Figure 5: the clustering of passengers according to *Regular* and *Going home* goal with a predominance of the “*Arrival*” strategy for this combination.

“Platform” strategy - Further subcategories

Concerning “*Platform*” strategy, as said before, due to its richness, further analyses include three subcategories “*Entries*”, “*Less crowded*” and “*Others*”.

Table 2: relationships between *Entries*, *Less crowded* or *Others* categories of motives with passengers’ properties.

	Entries	Less crowded	Others
Age	37.7 (sd:13)	40.5 (sd:11.3)	37.4 (sd:13.9)
Female	0.53	0.45	0.42
Regular	0.80	0.90	0.90
Going home	0.94	0.93	0.94
Travelling alone	1.00	0.98	0.89
N	71	70	50

“*Entries*” users are those who choose their position by minimizing the distance from the station entrance to the platform. “*Less crowded*” users are passengers who choose their position according to the decreasing density of people on the platform. Finally, “*Others*” category includes a variety of reasons of choice of location (because of platform seats, support, snack machine, more pleasant place...).

Table 2 shows that being regular passengers made providing a smallest number of “*Entries*” motives and being in a group of passengers a smallest number of “*Others*” motives.

“Entries” strategy:

Results show that “*Entries*” users tend to include more *Irregular* users and this motive was related with shortest *Times* duration compared to the other strategies. This could be explained by the fact that *Irregular* users do not have established habits about what they could gain (other than the distance at the entrance). Also, shortest *Times* of the journey may discourage users to make the effort to look for a strategy as their journey duration is not too long.

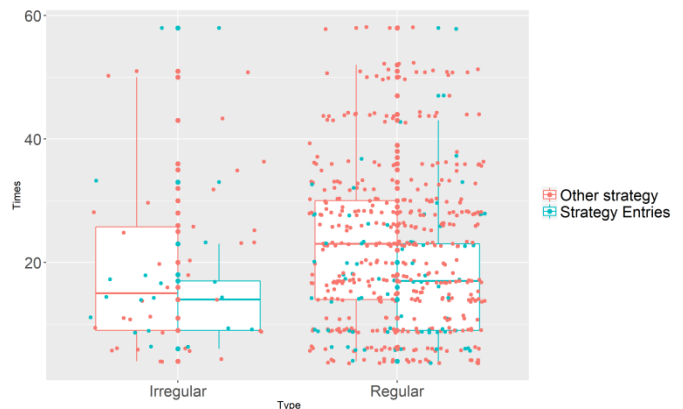


Figure 6: the relations between *Regular* and *Irregular* users with “*Entries*” motives.

“Less crowded” strategy:

According to our analyses, “*Less crowded*” strategy couldn’t be linked with our individual and trip variables.

Environment and contextual characteristics

The above strategies can be investigated according to properties of the environment such as entries and exits destination positions, axle load and density on the platform. To do so, we compute for each zone,

- the number of persons entering to the platform (*Entrants*),
- the number of persons on the platform (*Density*);
- the exits positions of the top ten destinations and their corresponding location in the station platform (*Exits*),
- the number of boarding and alighting passengers (*In / Out*),
- the axle load (*Load*)

We essentially identify a relation between “*Entries*” strategy and the number of *entrances* on the platform. This is quite expected since this strategy includes behavior of clustering around entrances. Also “*Less Crowded*” strategy seems to be related to density: as this last decreases, strategy “*Less crowded*” increases.

We do not obtain a significant relation effect between destination exits position and “*Arrival*” strategy. However, we computed these data for the exits location of the most visited destinations from BFM platform. Descriptively, users’ locations show normal distributions around these exits positions and the neighboring zones.

We do not obtain significant effect of relations between *axle load* and “*Train*” strategy; we thought that prior information about how people are occupying wagons would have impacted vehicle selection. In our case, it appears a slight effect of the number of boarding and alighting users. Maybe the perceived number of other boarding and alighting passengers influences one own judgment: more alighting persons from the train implies more free seats. More boarding implies less chance to obtain a free seat. Lastly, we evaluate odds values (Table 3).

Table 3: most important odds values for each variable model

Variable	Platform		Exits	Train	
	Nb Entrants	Density		In	Out
Entries	1.003				
Less crowded		0.98			
Arrival			1.14		
Train				0.95	1.14

Modeling and variable choice

We selected relation between variables and behaviors and will try to refine our models. Because our Multiple Linear Regression involves a few predictor variables, we use Akaike’s Information Criterion (AIC; see, e.g., Akaike, 1973, 1974, 1987) to determine the best model to use (Table 4). Dependants’ Variables (DV) are strategies that we try to explain through Independent Variable (ID) which are individual, trip and contextual variables:

DV: *Arrivals, Train, Platform* including *Entries, Less crowded* and *Others*.

ID: *Age, Gender, Type, Motif, Volume, Times, Nb Entrants, Density, Load, In, Out*

We notice that the subcategories “*Platform*” could in some way build the “*Platform*” best fitting model, except for the “*times*” variables that appear only for “*Entries*” subcategory. It appears to have, at the same time, a cumulative and neutralizing effect.

Table 4: Lowest AIC values determines the reported Best fit models for each strategy

Variable	Best fitting model	
	Individual / Trip	Contextual
Platform	$\sim \beta_0$ + Volume + Type	
Entries	$\sim \beta_0$ + Times + Type	+ Entries
Less crowded	$\sim \beta_0$	+ Density
Others	$\sim \beta_0$ + Volume	
Train	$\sim \beta_0$ + Age + Times.Int	+ In + Out
Arrivals	$\sim \beta_0$ + Motif + Type	

Times.Int is discretization of *times* in 10 minutes interval

For more precision, we calculate the odds values for each variable model. The odd value expresses the chance to see appearing a strategy per one unit variable.

For example, the odds of the “*Platform*” strategy are multiplied for each additional variable: by 2.4 being an irregular passenger (I) and by 8.5 for traveling in a group (G).

Table 5: most important odds values for each variable

Variable	Volume	Type	Motif	Gender	Time	Age
Platform	8.5(G)	2.4(I)				
Entries		1.94(I)			1.04	
L-crowded						
Others	13.05(G)					
Train						1.02
Arrival		1.9(R)	2.5(H)			
No reason			16.42	3.9(M)	1.04	1.03

Conclusion

In short, we obtain *Intentional* choices that encompass *Platform, Train, Arrival* strategies, which have been investigated in detail next to individual-specific, trip related and contextual variables. We established, through GLM, six models related to these strategies and the cited variables (table 4), describing sufficiently waiting strategies on platform.

Our findings correspond to those of Kim et al. (2014) work: they reported users aiming to minimize distance at arrival (69.7%), minimizing distance at entrances (16.6%) and maximize its comfort during journey (13.5%).

In our case, the crowd avoidance component appears clearly and would be investigated next with proxemic studies: how people behave in space (Hall, 1966). These findings will be incorporated onto the SpirOps simulator (see the SpirOps demo) with respective dependencies. Thus, we will implements causalities of motivation instead of observed behavior. The simulation will allows us to calibrate and valid the model and to improve its robustness and to ensure obtaining a prospective and predictive tool.

The next steps will be about understanding transversal distribution. At this stage, we are able to create a reliable simulation of waiting platform strategies. As we are interested on crowd dynamic and its corresponding effect on dwell-time, this work requires additional investigations on other contexts: behaviors inside the train and behaviors on Platform Train-Interface (PTI) while transferring. Many studies are already focussing on these topics (Berkovich et al., 2013; Lau, 2005; Hirsch & Thompson, 2011).

Acknowledgments

The authors would like to thank ANRT (*Association Nationale de la Recherche et de la Technologie*) for the project funding which takes place at *SNCF Réseau*.

References

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds.), *Proceedings of the Second International Symposium on Information Theory* (pp. 267-281). Budapest: Akademiai Kiado.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.
- Akaike, H. (1987). Factor analysis and AIC. In *Selected Papers of Hirotugu Akaike* (pp. 371-386). Springer, New York, NY.
- Berkovich, A., Lu, A., Levine, B., & Reddy, A. (2013). Observed customer seating and standing behavior and seat preferences on board subway cars in New York City. *Transportation Research Record: Journal of the Transportation Research Board*, (2353), 33-46.
- Boero, R. (2015). *Behavioral computational social science*. John Wiley & Sons.
- Buendia, A. (2003). *Une architecture pour la conception de modèles décisionnels: application aux créatures virtuelles dans les jeux vidéos* (Doctoral dissertation, Paris 11).
- Constant, T., Buendia, A., Rolland, C., & Natkin, S. (2015). Enjeux et problématiques de conception d'un jeu sérieux pour la prise de décision. *Ingénierie des Systèmes d'Information*, 20(1), 107-131.
- de Lavalette, B. C., Tijus, C., Poitrenaud, S., Leproux, C., Bergeron, J., & Thouez, J. P. (2009). Pedestrian crossing decision-making: A situational and behavioral approach. *Safety science*, 47(9), 1248-1253.
- Fernández, R., Valencia, A., & Seriani, S. (2015). On passenger saturation flow in public transport doors. *Transportation Research Part A: Policy and Practice*, 78, 102-112.
- Helbing, D., & Molnar, P. (1995). Social force model for pedestrian dynamics. *Physical review E*, 51(5), 4282.
- Hall, E. T. (1966). *The hidden dimension*. Garden City, New York: Doubleday & Company.
- Hirsch, L., & Thompson, K. (2011). *Tarzan travellers: Australian rail passenger perspectives of the design of handholds in carriages* (Doctoral dissertation, Human Factors and Ergonomics Society of Australia).
- Henderson, L. F. (1974) On the Fluid Mechanics of Human Crowd Motion. *Transpn. Res.* 8, 509. 5.
- Hutton, B. (2013). *Planning sustainable transport*. Routledge. London : Routledge.
- Kim, H., Kwon, S., Wu, S. K., & Sohn, K. (2014). Why do passengers choose a specific car of a metro train during the morning peak hours?. *Transportation research part A: policy and practice*, 61, 249-258.
- Lau, S. (2005). *Evaluating Interior and Door Configurations of rail vehicles by using variable loading densities*. Transport Research Record: Journal of the transportation Research Board, No 1927.
- Macal, C. M., & North, M. J. (2010). Tutorial on agent-based modelling and simulation. *Journal of simulation*, 4(3), 151-162.
- McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological review*, 96(4), 690-702.
- McCullagh, P. och Nelder, JA (1989). *Generalized Linear Models. Monographs on Statistics and Applied Probability. Chapman & Hall.*
- Puong, A. (2000). Dwell time model and analysis for the MBTA red line. *Massachusetts Institute of Technology Research Memo*.
- Spier, E., & McFarland, D. (1997). Possibly optimal decision-making under self-sufficiency and autonomy. *Journal of theoretical biology*, 189(3), 317-331.
- SpirOps
demo, http://www.spirops.com/SAMIE/SNCF/Articles/SpirOps_Simulation_BNF.mp4
- Yamada, S., Terabe, S., & Kasai, M. (2014). *Safety performance measures for railway stations* (No. 14-1868).
- Zhang, S., Luo, Y., Chen, B., & Bao, X. (2009). Progress of Simulation Studies of Pedestrian Traffic. In *International Conference on Transportation Engineering 2009* (pp. 2596-2601).