Proxymix: Influence of Spatial Configuration on Human Collaboration through Agent-based Visualization

Nicolas Ayoub[a], Cristian Jara-Figueroa[a], and Arnaud Grignard[a,*]

[a]MIT Media Lab, Massachusetts Institute of Technology
[*]E-mail: agrignard@media.mit.edu

November 2019

Abstract

This study proposes the use of agent-based simulation as an alternative to *space syntax*, a common technique in architecture, to reveal how architectural design can influence scientific collaboration. Using the MIT Media Lab building as a case study, we use Gama-platform to implement a parsimonious agent-based model of researchers' daily routine as they move inside the space. We find that the simulated collaboration network predicts the ground truth collaboration inferred from the Media Lab project database, even after controlling for institutional barriers such as the research lab researchers belong to. Our results highlight that agent-based simulation can be used to construct flexible indicators from architectural blueprints that reveal important characteristics of people's interaction inside the space.

1 Introduction

What is the influence of building configuration on social life? For decades, research on space syntax has tackled this question by developing quantitative descriptions of the built environment aimed at making their implied social logic more transparent [1, 2]. The premise of this research program is that the social aspect of space can be inferred from the relationships between socially relevant aspects. Most space syntax methods rely on converting physical layouts into spatial networks, using metrics such as walking distance, centrality, or visibility. From the early work by Thomas Allen [3], one area of particular interest has been the impact of the built environment on scientific collaboration. Recent work has shifted towards metrics based on space syntax that are more sophisticated than walking distance. For example, [4] shows that the overlap between the entry and exit routes from two offices correlates with scientific collaboration between the researchers located in those offices.

Yet, the purely spatial approach that most space syntax research follows can sometimes be inconsistent, too simple, and rigid for particular applications [5]. These and other critiques, especially targeted at the standard method of axial analysis, have lead researchers to consider alternatives based on complex systems [6] and computer simulations [7], especially in urban planning. Yet, in architecture, due to the lack of accessible tools to develop and communicate models, the lack of consensus on a standard, the lack of specialized researchers (there are more computer scientists in urban studies that in architecture), and the relative skepticism on the usefulness of these tools for design purposes, computational modelling remains widely unused[1].

Proxymix proposes a more flexible alternative; an agent-based visualization tool to understand the influence that spatial configuration has on human collaboration. This agent-based model at the architectural scale enables more optimal uses of space by highlighting the implied consequences on human behavior during the design process. Departing from traditional space syntax methods, Proxymix presents a more transparent approach in which the assumptions about how individuals use the space are formalized in an agent-based model. Building on the research of Edward T. Hall and Thomas Allen [8], we use Proxymix to investigate the concept of proxemics, which examines "cross-cultural communication processes and the design of built environments" [9], by analyzing the case of the MIT Media Lab. The Media Lab provides an interesting case study that slightly departs from previous methods that have relied on measuring scientific collaboration by looking at co-authorship on published academic research. The "projects" that we use as a signal of collaboration are often very early-stage and driven mainly by researchers.

While urban studies have already adopted agent-based models as part of their toolbox, this is not true for architects. Most of the computational tools used by architects circle around the digitization of the traditional drawing board. Computer-aided design software (CAD)

[1]With the notable exception of Building Information Modelling (BIM) methodologies.

such as Autodesk's AutoCAD are widely used to draw static representations of the built environment but are limited when dealing with complex hierarchical dynamics. In other words, CAD software cannot deal with either the temporal dimension or with any form of organizational hierarchy. Notable strides forward have been made within the Building Information Modelling (BIM) community. While dynamic, those methods have yet to integrate information about the relational structure of space. ABM has been used in architecture for occupant behaviour models [10, 11, 12] for building performance simulations such as heating, ventilation, and energy consumption [13], as well as evacuation scenarios in emergency situations [14]. Processes of people collaboration in human-building interaction have yet to be extensively explored. Proxymix takes a step in this direction, by visualizing and simulating those behavioral patterns in the built environment.

2 Methodology: Coupling spaces and behavior

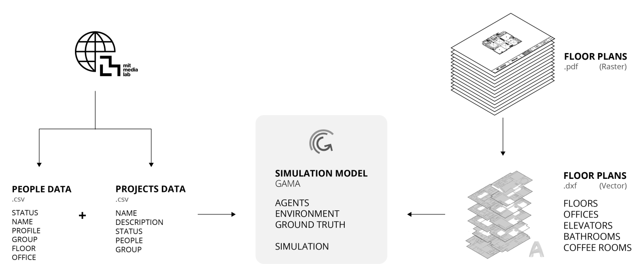

Figure 1: Description of the methodology used for Proxymix. The model relies on data from the Media Lab directory that contains information about office assignment, combined with the building blueprints. The project database is used to infer a ground-truth collaboration network as a way to validate the model.

This study follows a three step approach to build a spatially aware model of research collaboration inside the MIT Media Lab, described in Figure 1. First, we collect data on both the building blueprints and on the office assignment for all Media Lab researchers. This data also included the projects each researcher has developed. Second, we design and implement a model on Gama [16] that simulates a plausible daily routine for each researcher. Finally, we compare the network of interaction implied by our agent-based model with the ground truth information about collaborations inferred from the Media Lab's project database.

While architects use CAD software to draft floor plans, the final output is often rasterized for printing or digital publishing purposes. As such, the raster file becomes merely a visual representation of space, discarding its inherent vector attributes. The issue of raster-to-vector for-

Overview	Based on a daily pattern activity model, a synthetic population is created. The model describes 2 different profiles having a slightly different behaviour.
Design	Researcher profile: 8am-10am: Arrive to your office from an elevator 8am-12pm: Grab a random coffee before lunch 12pm-1pm: Collective lunch 1pm-6pm: Go back to your office + random coffee + random bathroom 6pm-8pm: Leave the building, go home Student profile: 8am-10am: Arrive to your office from an elevator 8am-12pm: Go to morning class 12pm-1pm: Collective lunch 1pm-2pm: Go back to your office 2pm-4pm: Leave the building, external class
Details	The model relies on the digitized building blue-prints and the office assignment for all researchers in the building.

Table 1: ODD description of the WLCB Model [15].

matting is a longstanding one in the architectural realm, and has yet to find a permanent workaround. Here we recovered the "lost" vector data by retracing the floor plans from scratch on AutoCAD.

The MIT Media Lab is a research lab inside the Massachusetts Institute of Technology that is funded mainly by corporate money. Because of this, it maintains a project database that is updates twice a year to communicate progress to "member companies". The database consists of multiple projects, each associated to a group of researchers through their Media Lab user name. This association is used to match each researcher with the Media Lab directory that assigns them to an office or an open space inside the building. Because the Media Lab is still an academic research institution, most research collaborations are driven by the researchers themselves, with little top-down assignment. Creating a Media Lab project is relatively cost-less, which makes these data interesting to study scientific collaboration since projects can either be mature research projects or very early-stage and student-driven. This stands in contrast with traditional approach relying on publishing academic papers.

We use data for all projects created as of 2015 and all active researchers as of June 2019.

Agent-based modelling is a class of computational models that simulates the actions and interactions of autonomous agents in order to generate global effects. ABM is used in many different fields such as biology, network

Figure 2: Proxymix; visualizing the impact of space on scientific collaboration using the MIT Media Lab as a case study. The displayed network is the collaboration network implied by the model, using a threshold of 2 meters. Every agent is placed in their office according to the information available in the Media Lab directory and their daily routine is simulated according to the model described in Table 1. The information on the right-hand side panel is obtained from the Media Lab project database.

theory, economics, and social science. The interest in agent-based models is mainly due to a few recent advances: (i) the improvement of computational capacities and rapid growth in data availability, (ii) the introduction new visualisation modes like augmented, virtual, and mixed reality, (iii) new ways of interacting with simulations, and (iv) new ways of constructing and coupling models with more credible behavior and a better integration.

Here we use Gama [16], a spatially explicit agent-based simulation platform used for many applications. Since the latest version of GAMA released in 2018, AutoCAD format (dxf) are easily integrated in the platform. The WLCB Model is described in Table 1.

Comparing the results of our model with the ground truth collaboration data amounts to comparing a weighted graph with a non-weighted one. The model simulates the average distance between every two pairs of researchers, which needs to be compared to a binary variable indicating whether the two researchers in the pair have collaborated in the past four years or not. To do this, we resort to a receiver operating characteristic (ROC) curve that compares the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings for the average distance. As a measure of goodness-of-fit we use the area under the ROC curve (AUC). As a reference, the ROC curve of a random model will be a diagonal line (the TPR and FPR are the same) and its AUC will be 50%.

The comparison between the simulated and ground truth graphs needs to take into account the fact that there are institutional barriers to collaboration within the Media Lab. Namely, the Media Lab is structured into research "groups" that may or may not collaborate between them. There are not explicit policies that prevent researchers from collaborating across groups, but in practice we do observe much more collaboration within groups than between groups. Moreover, while researchers in the same group tend to occupy the same physical space, there are some groups that are spread across the building and some that share the space with other groups.

3 Results: Comparing ground truth and simulated graphs

Figure 2 shows the design of the Proxymix, a visualization tool for scientific collaboration. The connections are drawn according to the simulated graph with a threshold distance of 2 meters. The data on the right-hand side panel is obtained from the Media Lab project database.

Figure 3-A shows the ROC curves for our agent-based model (blue) and two other baseline models. Given the Media Lab organizational structure, we compare our

model the one that assumes collaboration only happens within groups. When comparing the yellow and blue curves in Figure 3-A we see that our model is as good as assuming that collaboration happens within the same research group.

Table 2 formally compares the relevance of the average distance obtained from our agent based model with the same-group predictor by using a logistic regression model. These results imply that, even after accounting for institutional barriers, an additional 10 meters of separation translates into a decrease of 4.5% in the chance of collaborating. To put this number in perspective, let us note that the unconditioned probability of collaborating is of 4.9%, which means that those additional 10 meters translate to a 90% decrease over the average. Figure 3-B shows the ROC curves for the three models from Table 2.

	Collaboration (binary variable)		
	(1)	(2)	(3)
Distance	-0.0445***		-0.0123***
(ABM)	(0.00347)		(0.00278)
Same group		0.136***	0.0969***
		(0.00567)	(0.00834)
Overlap time	0.171***	0.121***	0.129***
at ML	(0.0203)	(0.0169)	(0.0172)
N	3916	3916	3916
AUC	0.913	0.916	0.934

Marginal effects; Standard errors in parentheses

(d) for discrete change of dummy variable from 0 to 1

Distance measured in decimeters.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: Logistic regression predicting collaboration for each pair between 89 researchers in our data (3916 pairs). Model (1) shows that each additional 10 meters of separation in the model translates into a decrease in 4.5% in the chance of collaborating. The unconditioned probability of collaboration is $\sim$ 4.9%. After controlling for institutional barriers in model (3) (whether researchers are in the same group or not), the effect of the modelled distance decreases significantly: an increase in 10 meters of separation translates into a decrease in 1.2% in the change of collaborating. The traditional way of modelling the connection between distance and collaboration is by log-transforming the distance (collaboration decays exponentially with distance [3]). Here we have chosen to use a linear model because distances are relatively short and because we want to prioritize interpretability of the coefficients.

4 Conclusion

Given the increasing predominance of teams in the production of knowledge [17], it is increasingly more important to understand the factors that determine scientific collaboration. For decades, researchers have known that characteristics of the built environment play a key role [4, 3]. Yet, the environment does not directly shape collaboration but it rather shapes people's behavior, which leads to changes in collaboration patterns. Here we propose a methodology that explicitly assumes the way the built environment affects behavior, and models the collaboration as an outcome of this behavior. While more work is needed to refine these class of models [18], we believe that our contribution highlights the need for more explicit spatial behavioral models.

References

[1] Bill Hillier, Julienne Hanson, and John Peponis. What do we mean by building function? E & FN Spon Ltd, 1984.

[2] Sonit Bafna. Space syntax: A brief introduction to its logic and analytical techniques. *Environment and behavior*, 35(1):17–29, 2003.

[3] Thomas J Allen. Communication networks in r & d laboratories. *R&D Management*, 1(1):14–21, 1970.

[4] Felichism Kabo, Yongha Hwang, Margaret Levenstein, and Jason Owen-Smith. Shared paths to the lab: A sociospatial network analysis of collaboration. *Environment and Behavior*, 47(1):57–84, 2015.

[5] Carlo Ratti. Space syntax: some inconsistencies. *Environment and Planning B: Planning and Design*, 31(4):487–499, 2004.

[6] Andres Sevtsuk and Michael Mekonnen. Urban network analysis. *Revue internationale de géomatique–n*, 287:305, 2012.

[7] Arnaud Grignard, Luis Alonso, Patrick Taillandier, Benoit Gaudou, Tri Nguyen-Huu, Wolfgang Gruel, and Kent Larson. The impact of new mobility modes on a city: A generic approach using abm. In *International Conference on Complex Systems*, pages 272–280. Springer, 2018.

[8] Thomas Allen and Gunter Henn. *The organization and architecture of innovation.* Routledge, 2007.

[9] Nina Brown. Edward t. hall, proxemic theory, 1966. csiss classics. 2001.

[10] Wout Parys, Dirk Saelens, and Hugo Hens. Coupling of dynamic building simulation with stochastic modelling of occupant behaviour in offices–a review-based integrated methodology. *Journal of Building Performance Simulation*, 4(4):339–358, 2011.

[11] Davide Schaumann, Seonghyeon Moon, Muhammad Usman, Rhys Goldstein, Simon Breslav, Azam

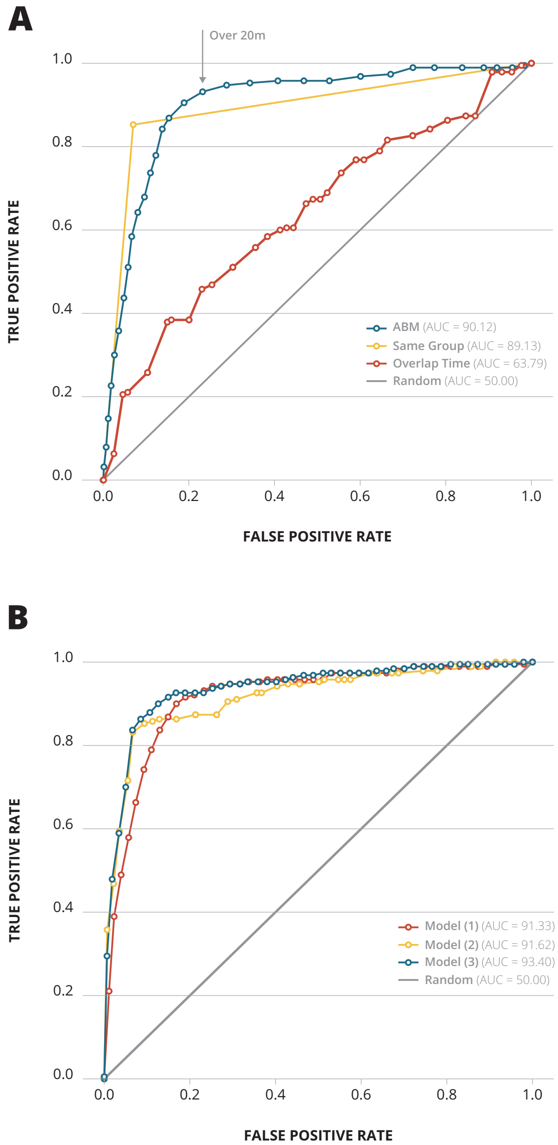

Figure 3: ROC (Receiver operating characteristic) curves comparing different models predicting collaboration of each pair of researchers. ROC curves are a standard way of comparing the predictive power of models that predict a binary outcome by plotting the true positive rate vs. the false positive rate for different thresholds of the model. The area under the curve (AUC) is often used as a goodness-of-fit parameters, where an AUC of 50 corresponds to the diagonal line of the random predictor. **A** shows the ROC curves for our agent-based simulation, a model that predicts collaboration using a dummy variable for belonging to the same group, and a model that predicts collaboration based on the amount of time in months that both researchers overlapped at the Media Lab. **B** shows the ROC curves for the three models presented in Table 2. While the same-group predictor already does a very good job at predicting collaboration, highlighting the importance of institutional barriers, the ABM approach proposed here still adds some predictive power for mid distances (over 20 meters).

Khan, Petros Faloutsos24, and Mubbasir Kapadia. Toward a multi-level and multi-paradigm platform for building occupant simulation.

[12] Davide Schaumann, Simon Breslav, Rhys Goldstein, Azam Khan, and Yehuda E Kalay. Simulating use scenarios in hospitals using multi-agent narratives. *Journal of Building Performance Simulation*, 10(5-6):636–652, 2017.

[13] Da Yan, William O'Brien, Tianzhen Hong, Xiaohang Feng, H Burak Gunay, Farhang Tahmasebi, and Ardeshir Mahdavi. Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy and Buildings*, 107:264–278, 2015.

[14] Shawn Singh, Mubbasir Kapadia, Billy Hewlett, Glenn Reinman, and Petros Faloutsos. A modular framework for adaptive agent-based steering. In *Symposium on Interactive 3D Graphics and Games*, pages PAGE–9. ACM, 2011.

[15] Volker Grimm, Uta Berger, Donald L DeAngelis, J Gary Polhill, Jarl Giske, and Steven F Railsback. The odd protocol: a review and first update. *Ecological modelling*, 221(23):2760–2768, 2010.

[16] Arnaud Grignard, Patrick Taillandier, Benoit Gaudou, Duc An Vo, Nghi Quang Huynh, and Alexis Drogoul. Gama 1.6: Advancing the art of complex agent-based modeling and simulation. In *International Conference on Principles and Practice of Multi-Agent Systems*, pages 117–131. Springer, 2013.

[17] Stefan Wuchty, Benjamin F Jones, and Brian Uzzi. The increasing dominance of teams in production of knowledge. *Science*, 316(5827):1036–1039, 2007.

[18] Mathieu Bourgais, Patrick Taillandier, Laurent Vercouter, and Carole Adam. Emotion modeling in social simulation: a survey. *Journal of Artificial Societies and Social Simulation*, 2018.