# Essays on Analytic Methods Applicable to the Micro-Geography of the Workplace

by

**Yongha Hwang**

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Architecture)
in the University of Michigan
2014

Doctoral Committee:

Professor Jean D. Wineman, Chair
Professor Gerald F. Davis
Professor Linda N. Groat
Professor John Peponis, Georgia Institute of Technology

To my wife, Woo Jong

# ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. Jean D. Wineman for her guidance in the preparation of this dissertation, and for her support throughout the course of my doctorate research. I would also like to thank all the members of my committee, Dr. Gerald F. Davis, Dr. Linda N. Groat, and Dr. John Peponis for their time and constructive criticism.

I am indebted to my former advisor, Dr. Jaepil Choi at Seoul National University. I also thank my co-authors, Felichism Kabo, Margaret Levenstein, and Jason Owen-Smith. I further wish to express my thanks to former colleagues, Sung Hoon Yoon, Jinwoo Moon, Sung Kwon Jung, and Youngchul Kim.

My thanks for emotional and practical supports in many aspects also go to the friends in the Parc Pointes, Sungsoon Cho, Hyunchul Park, Sangwon Seo, Yongjun Park, and Hyoun Kyu Cho.

Finally, I would like to thank my wife Woo Jong for her support and patience.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF EQUATIONS

# LIST OF ABBREVIATIONS

| | |
|---|---|
| SPSN | Spatial Projection of the Social Network |
| SSB | Sociospatial Betweenness |
| SPB | Spatial Betweenness |
| VGA | Visibility Graph Analysis |
| VOS | Visualization of Similarities |
| NCVG | Network Communities in the Visibility Graph |

# ABSTRACT

In spite of growing interest in the physical environment's role in better communication and collaboration in knowledge-intensive organizations, far too little attention has been paid to quantitative methods for describing and analyzing micro-geography of the workplace. Three essays in this dissertation explores novel methods for describing a spatial layout and analyzing its effect on organizational communications.

The first essay's main question is how concentration of movement fosters diverse communication in the space. We articulate the concept of confluence and propose a new metric, *sociospatial betweenness* to measure the confluence of a space. Sociospatial betweenness of a space was found to be positively associated with the diversity of communication partners among a group of professionals in a manufacturing company; in contrast, traditional spatial betweenness did not show such an association.

The second essay addresses how exposure between members of a dyad increases the chance of research collaboration. The essay proposes and develops a novel metric, *zone overlap*, measuring exposure, the likelihood of mutual encounter between two people, based on the location of one's workstation and commonly used facilities. We collected administrative data on a sample of research scientists working at two biomedical research buildings with different layouts. We found that increasing path overlap is associated with

increases in collaborations in both buildings. In contrast, traditional metrics such as walking distance and straight-line distance influence outcome measures in only one of the research buildings.

The third essay introduces a novel approach for subspace decomposition that can be used for the two new metrics, zone overlap and sociospatial betweenness, proposed in the two previous essays. Although spatial decomposition is one of the essential processes for the analysis of building layout, no new rigorous decomposition has been proposed for more than a decade until this study. We demonstrated that the new method successfully addresses the problems of traditional methods. The essay introduced the modularity function as a quality function to evaluate the goodness of spatial decomposition. Previous decomposition methods so far have rarely paid attention to the evaluation of decomposition.

# CHAPTER 1
# Introduction

It is an interesting trend that contemporary high tech companies providing services that connect remotely located people are eager to locate their employees physically together. In 2013, Yahoo's CEO Marissa Mayer sent a controversial memo to Yahoo's employees about a new policy banning working from home based on the belief that innovation comes from physically being together (Miller and Rampell 2013). Google, the former employer of Marissa Mayer, recently announced the plan of its new campus where "no employee in the 1.1-million-square-foot complex will be more than a two-and-a-half-minute walk from any other" to maximize serendipitous encounters among its employees (Goldberger 2013). Facebook, a big rival of Google, broke ground on its second campus designed by Frank Gehry in 2013. Facebook's 2,800 employees will be accommodated in a large, "one room building" intended for better communication and collaboration (Brown 2012). Zappos, one of the largest online shoe retailer, closed off a convenient skywalk directly connecting to a parking structure from its headquarter. Tony Hsieh, the CEO of Zappos, explained that he expected employees would collide more with their

colleagues by forcing them to walk through the lobby (Silverman 2013). Steve Jobs once adopted a radical approach for the Pixar studios. He insisted removing all other restrooms but two in the central space even though some people had to walk 15 minutes to go to the restroom after his radical approach. By removing other restrooms, Jobs could force people to meet and mingle with each other around the two remaining restrooms (Lehrer 2012).

## 1.1. From Proximity to Exposure

Such attitudes of high-tech companies regarding their physical environments is based on the expectation that the physical environment of workplaces affects communication among employees. There is academic research behind such entrepreneurial expectations.

The famous Allen study (1977) demonstrated how the frequency of communication among professionals diminishes as proximity between them increases. For a total 512 researchers in seven R&D laboratories, Allen investigated the relationship between walking distance and the probability of communication about 'technical or scientific matters'. The result is the famous 'Allen curve', an L-shaped curve showing the probability of communication falls rapidly as the separation distance increases. To control the effect of organizational structure, he separated the curve into two groups: pairs with organizational bonds and pairs without organizational bonds. Both of the curves representing each group also showed L-shaped curves.

In some studies, distance is defined with ordered categories. Kraut and colleagues (1988) showed that physical distance between scientists is related to the formation of research collaboration in a large industrial R&D laboratory. In their study, distance between a pair

of researchers was measured with ordered categories such as ‘same corridor’, ‘same floor’, ‘different floor’, and ‘different building’. Such a tendency was also observed when department affiliation is controlled. Another observational study on the interactions among a group of 12 office workers demonstrated that proximity was the most important factor to the frequency of interactions (Gullahorn 1952). The workers were seated in three rows of four people each. Through two week of observation, the author found that within-row interaction is far more frequent than across-row interaction; interaction between adjacent rows is far greater than interaction between separated rows; within each row, interaction frequency is higher when workers are seated closely.

Frequent interactions have also been found to enhance shared culture. A recent study on a prediction market in a global corporation illustrated that employees tend to share the same position on future predictions when they are closely located after controlling department or team affiliation, existing social connection, demographic similarity, work history (Cowgill, Wolfers, and Zitzewitz 2009). For over 140,000 dyads within and across the four campuses of the company, the proximity of dyads at several levels were measured: same city, proximity within city (inverse of the straight line distance between buildings), same building, same floor, proximity on floor (inverse of the straight line distance between offices within a floor), and same office. It was found that proximity within city, same floor, proximity on floor, and same office were relevant predictors of shared views.

A study on academic faculty’s collaboration patterns revealed that faculty members collaborate more when they are closely located (Wineman, Kabo, and Davis 2008). The authors collected CV’s of 82 faculty members in a professional school who had co-

authored with another faculty member in the school. For each dyad among all possible dyads from 82 faculty members, they investigated the effect of distance measured by the number of 'turns' to the likelihood of co-authoring after controlling the effect of departmental affiliation. They found that the closer faculty offices were located, the more likely the faculty members were to collaborate. Another study on four organizations sized from 63 to 120 reaffirmed physical proximity's association to a social tie formation (Sailer and McCulloh 2012). The authors illustrated that walking distance between two people's offices was a good predictor of self-reported interaction frequency between the two people.

Festinger and colleague's seminal work (1950) on how spatial structure affects social tie formation had often been neglected in recent space-communication literature. However, it is worth revisiting as it gives an alternative perspective on the effect of spatial structure. The authors analyzed the effect of space on the formation of social relationships based on the analysis of friendship formation among freshmen in a student housing complex. They found that the proximity of student rooms was a strong predictor of friendship formation. It is notable that the authors point out that the effect of space cannot be reduced to the effect of distance. The authors claim that the specific patterns of movement generated by spatial structure should be considered in the analysis of the effect of space. For example, even though distance between A and B is equal to B and C, B is more likely to become a friend of A if B should pass in front of A's door to use a stair to the ground level. What is more important here is how likely two people are to see each other, than how close in distance they are.

This tendency that people interact more when they see each other more frequently is usually explained by mere-exposure effect (Kraut, Egido, and Galegher 1988). Mere-exposure effect is a psychological concept describing people's tendency to respond positively to a stimulus simply because they are exposed to the stimulus repeatedly (Zajonc 1968). When two people see each other more frequently, they tend to rate each other more positively and hence are more likely to initiate a conversation. Explaining the effect of distance to the frequent interaction with mere-exposure effect leads us to see distance as a proxy of the likelihood of exposure. In other words, distance is a predictor of interaction frequency because closely located people are more likely to see each other than remotely located people.

Then we may want to find other possible proxies of the likelihood of mutual exposure. Monge and Kirste (1980) criticized using physical distance as a proxy of the likelihood of mutual exposure, arguing that it assumes people are fixed in one location and the location is not changed over time. They proposed to use a more direct metric measuring the likelihood of encounter than physical distance. The authors asked 75 respondents to indicate the number of hours each respondent spent at each room in the workplace in an 'average' week. For all 2775 dyads from 75 respondents, the authors calculated the joint probability of co-occupancy over all rooms. Then the joint probability of a dyad is compared to the self-reported number of minutes the dyad communicated. They found a significant association between the two variables.

In this study, we propose *exposure,* the likelihood of mutual encounter between two people. It is a new spatial property related to interaction behavior at the workplace. As Monge and Kirste's study suggests, the traditional concept of proximity does not

effectively represent the spatial relationship between two people in terms of the possibility of mutual communication. For example, if there are two employees both of whom are heavy coffee-drinkers, their *exposure* would increase when a nice coffee bar opens in the floor even though their physical distance remains unchanged. Although physical distance is a classic metric for exposure and has been extensively used because it is simple and intuitive, it fails to capture the subtle effect of spatial structure (See Chapter 3 for more detailed discussion). Monge and Kirste's joint probability of co-occupancy may be a more detailed metric of exposure. However, it is pricey because calculating the joint probability requires additional information on how much time a person spends in each room; whereas physical distance requires only a floorplan on which each person's primary location is marked. In other words, physical distance and joint probability are two extreme metrics for exposure: simple but coarse vs. detailed but expensive. In Chapter 3, we propose *zone overlap*, a new metric for exposure, balanced between the two extremes so that the new metric can capture the subtle effect of spatial structure with minimal additional information.

## 1.2. Accessibility, Visibility and Confluence

The studies reviewed in the previous section focus on interactions between a dyad among people in a workplace. The spatial element of interest in such studies is the path between two people in the dyad, and hence the spatial properties mainly used for the studies are the properties of the path like distance between two people, or the length of the path.

There are other types of spatial properties associated to interaction behaviors in workplace, which are used to answer questions like 'where are people likely to have

social interactions' instead of 'which pair of people are likely to have social interactions' as in the studies from the previous section. These are accessibility (Hillier and Penn 1991; Penn, Desyllas, and Vaughan 1999; Peponis 1985; Toker and Gray 2008) and visibility (Hatch 1987; Rashid et al. 2006; Steen and Markhede 2010; Stryker, Santoro, and Farris 2012; Appel-Meulenbroek 2010).

Studies that directly or indirectly used the accessibility of a space as a predictor of communication frequency are based on the core proposition that more communications tend to happen in a more accessible space than in a less accessible space. Penn and colleagues (1999) explain the core argument in this line of studies like follows. The more accessible a space is, the more people would be in the space. If there are more people in the space, more people are expected to encounter one another, and then to be recruited into interactions. Most of such studies have defined the accessibility of a space as the average distance from the focal space to all other spaces where distance is defined in a various ways such as the number of turns (Hillier and Penn 1991; Penn, Desyllas, and Vaughan 1999), the number of spaces (Toker and Gray 2008), and walking distance (Wineman et al. 2014).

One of the notable pioneering studies on the effect of accessibility to interaction in the workplace is the study by Hiller and Penn (1991). From this observational study on a floor occupied by a daily newspaper company, the authors partially demonstrated the core argument by showing that a space of higher accessibility also tends to have larger numbers of encounters ($r=0.83$, $p<0.001$). Later, from the analysis of an energy utility company, Penn and colleagues (1999) found additional bivariate associations between the accessibility of a space and the number of moving people in the space ($r=0.966$, $p<0.001$)

and also between the number of moving people in a space and the number of talking people (r=0.989, p=0.0001), both of which support the core argument.

The visibility of a space is another spatial property that has been used for analyzing the spatial structure's effect on communication behavior. It has long been a topic of debate whether high inter-visibility or less barriers in the layout would promote an organization's communication (Elsbach and Pratt 2007). There are two competing views on the relationship between inter-visibility and the amount of interaction. The first view is to support the association between high inter-visibility and more interaction (Oldham and Brass 1979; Brookes and Kaplan 1972; Allen and Gerstberger 1973; Peponis et al. 2007; Boutellier et al. 2008) claiming that inter-visibility provides visual information required for the initiation of informal interaction. The other view supports the opposite direction of the association (Boje 1971; Clearwater 1980; Hanson 1983; Sundstrom, Herbert, and Brown 1982) claiming that visual barriers reduce concerns of confidentiality and disturbance to coworkers, therefore promoting communication.

The unit of analysis of the majority of such studies investigating the effect of visibility has usually been a floor or a layout. These studies compare the level of interaction across different types of layouts, which is often operationalized with a categorical variable such as 'open', 'close/cell/traditional', or 'mixed'. There are a relatively small number of studies investigating the effect of visibility at more detailed level such as workstations or locations on a floor. Studies at the detailed level also yield mixed results.

Hatch's study (1987) on the effect of visual barriers surrounding one's workstation on work-related communication is one of the few empirical studies at the workstation level. For 99 employees in two high-technology firms, the author measured the number of

partitions and average height of partitions around each employee's workstation and collected work activity data using self-reported activity logs. Then the author analyzed the effect of these variables on interaction behaviors after controlling for type of tasks, demographic background, and position in one's organization. The author found a negative association between the visibility of one's workstation and the amount of a person's interaction activities. In contrast, Rashid and colleagues' (2006) observational study in four US federal offices supports the positive association between the degree of visibility and the amount of interaction. For each route segment on the floors, an observer recorded the number of face-to-face interactions and the number of people visible from the segment. Across all four offices, a space with more visible people tends to have more face-to-face interactions. Steen and Markhede (2010) also found a positive association between visibility and interaction in a study on a newspaper office. From the interaction data collected by self-reports, the authors found that a person whose workstation is seen from a larger area tends to have more interactions. Toker and Gray's study (2008) on six university research centers is one of the rare studies incorporating visibility and accessibility in a single study. From the analyses of self-reported activity logs from 114 scientists in the research centers, the authors found that both of the spatial properties of a scientist's workstation are positively associated with the frequency of organizational interactions of the scientist.

We introduce *confluence* as another spatial property expected to be associated with interaction behaviors in workplaces. Confluence is a property of a space describing how much movement flux would be concentrated in the space because of the spatial structure of the layout. The lobby of Zappos' headquarter or the restroom in the Pixar studios are

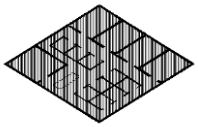
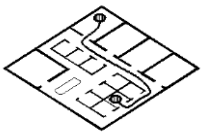
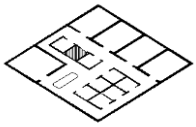
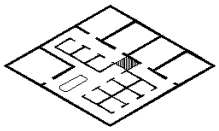
examples of the space with high confluence. Movement flux is concentrated in the lobby of Zappos as the lobby becomes the only connection to the parking building, or a bottleneck, when the skywalk is closed. The two restrooms in the Pixar studios are also spaces of high confluence not because they are bottleneck spaces but because they are highly 'attractive' spaces. In both cases, the concentration of movement to the lobby or the restrooms is not readily related to the accessibility or the visibility of the spaces. Thus we need to introduce confluence as an additional spatial property describing the degree of concentration of movements. In chapter 3, we propose and test a new metric *sociospatial betweenness* for measuring the confluence of a space.

## 1.3. Types of Space-Communication Studies

Table 1.1 summarizes the current geography of space-communication studies according to the kind of the spatial elements and social elements each study emphasizes. The table also delineates the contributions of this dissertation to the field.

The first type of study focuses on the overall *layout* of a floor. When a study in this field is focusing on the layout of a floor, the study's social element is usually a group of people occupying the floor. Most of the studies of this type have given emphasis to the openness, or the overall inter-visibility of a layout when they try to explain interaction behaviors in the layout. Thus a typical research question of the studies in this category looks like "Would people have more interactions in a more inter-visible layout?" The openness or the inter-visibility of a layout is usually quantified as a categorical variable according to layout type such as 'open' or 'closed'.

Table 1.1. Common types of space-communication studies

| Type | **Layout** | **Dyad** | **Location** | |
|---|---|---|---|---|
| |  |  |  |  |
| Spatial Element | layout of a floor | path between two people | individual office or workstation | subspace (usually in common areas) |
| Social Element | a group | a dyad | an individual | a group |
| Typical Question | Would a group have more interactions in a more inter-visible layout? | Would a dyad have more interactions if the two persons are closely located? | Would a person have more interactions if one's location is more accessible/visible? | Would a space accommodate more interactions if the space is more accessible/visible? |
| Explained Behavior | Interactions the group had in the layout | Interactions between the dyad<br>** Information on the both sides of interactions required* | Interactions the individual had | Interactions happened in the space<br>** Information on the locations of interactions required* |
| Spatial Property | • Openness | • Proximity<br>• ***Exposure*** | • Accessibility<br>• Visibility<br>• ***Confluence*** | |
| Metric | • Type of layout (open vs. cell) | • Same floor/building<br>• Number of turns<br>• Number of spaces<br>• Straight line distance<br>• Walking distance<br>• ***Zone Overlap*** | • Average distance to other spaces<br>• Average number of turns to other spaces<br>• Degree of enclosure<br>• Size of visible area<br>• Number of visible people<br>• ***Sociospatial Betweenness*** | |
| Exemplary Study | Hatch 1987<br>Toker and Gray 2008 | Penn et al. 1999<br>Rashid et al. 2006 | Allen 1977<br>Kabo et al. 2013 | Hanson 1978<br>Oldham and Brass 1979 |

The second type of study focuses on a *dyad* in the group of people in the workplace. When a dyad is a social element of analysis, the spatial property of interest usually becomes the proximity between two people in the dyad, or more generally, exposure

between the two people. A typical question becomes “Would a dyad have more interactions if the two persons are closely located?” Proximity has been usually quantified using various distance concepts such as number of turns, number of spaces, straight-line distance, and walking distance. This study discusses why exposure should be separated from proximity, and proposes a new metric for exposure. In contrast to other types of studies, a dyad-level study requires information on ‘with whom a person had interactions’ which requires more effort.

The main inquiry of the third type of study is how interaction behavior differs across *locations* in the layout according to the spatial properties of the locations. Thus the analysis is performed at the level of locations in the space such as a workstation, a room, a corridor, and a tessellated grid point. The third type of study has two subtypes whether a researcher’s emphasis is given more to a space or to a person. When the emphasis is more focused on a person, a typical question becomes “Would a person have more interactions if the he or she occupies a better location?” In contrast, when the emphasis is on to a space, a typical question becomes “Would a space accommodate more interactions from a group of people if the space occupies a better location?” It is noteworthy that this kind of question requires information on the locations of interaction, which is often hard to obtain. Two spatial properties, accessibility and visibility have been used to identify ‘a better location’ for frequent interactions. The accessibility of a space has been usually quantified with average distance to other spaces from the space. The visibility of a space has been typically quantified with the degree of enclosure of the space, the size of visible area from the space, or the number of visible people from the

space. This study proposes confluence as the third spatial property for measuring a space's advantage to interaction behaviors.

## 1.4. Space Decomposition

Recent space-communication studies often decompose a building layout into subspaces to apply a quantitative analytical method such as network analysis or visibility analysis. Three major approaches for decomposing a layout into subspaces have been used in space-communication studies: linear decomposition (Penn, Desyllas, and Vaughan 1999; Rashid et al. 2006; Peponis et al. 2007; Wineman, Kabo, and Davis 2008), subspace decomposition (Peponis 1985; Hillier and Penn 1991; Toker and Gray 2008), and grid decomposition(Peponis et al. 2007; Appel-Meulenbroek 2010; Steen and Markhede 2010).

Linear decomposition and subspace decomposition are usually used when accessibility is the spatial property of focus. Grid decomposition is used when detailed analysis on the visibility of a space is required. The two metrics, zone overlap and sociospatial betweenness, newly developed in this study utilize a subspace decomposition method in order to calculate the values of the two metrics. So far, little attention has been paid to the decomposition methods. This study proposes a novel subspace decomposition method so that researchers can use the new decomposition method for the two new metrics.

## 1.5. General Outline

In spite of growing interests in the physical environment's role in better communication and collaboration in knowledge-intensive organizations, far too little attention has been

paid to quantitative methods for describing and analyzing micro-geography of the workplace. Three essays in this dissertation explores novel methods for describing a spatial layout and analyzing its effect on organizational communications.

This dissertation is composed of 5 chapters: this introductory chapter, three independent essays, and a conclusion. The chapters are as follows: Chapter 2 is an essay on how exposure between a dyad increases the chance of research collaboration; Chapter 3 is an essay on how the confluence of a space fosters diverse communication in the space; Chapter 4 is an essay on a new space decomposition method based on a community detection algorithm; finally, Chapter 5 summarizes the findings from the three essays and provides conclusions.

In more detail, Chapter 2 articulates the concept of confluence and propose a new metric, *sociospatial betweenness,* to measure the confluence of a space. The new metric is obtained by projecting social network onto the spatial structure. This projection enables us to describe the spatial structure and social structure together; hence, it becomes possible to capture the subtle differences in socio-spatial condition due to the locations of people and the social ties among them. We analyze the association between sociospatial betweenness and the diversity of communication in a space using a dataset that included the locations of interactions among a group of professionals in a manufacturing company collected from location tracking sensors worn by each member of this group.

In Chapter 3, we propose and develop a novel metric, *zone overlap*, measuring exposure between two people based on the location of one's workstation and key facilities. While our focus was, in the previous chapter, on how the location of people and social ties differentiate socio-spatial condition, our focus is now on how the location of key facilities

differentiates the socio-spatial condition. The new metric based on the overlap of individual functional zones is a new way of describing 'proximity' between two people. We collected administrative data on a sample of research scientists working at two biomedical research buildings with different layouts during the period 2006-2010. The data includes floorplans, detailed space allocation for each scientist, and evidence of research collaboration such as grant applications. The regression analyses of collaboration rates are compared across the new metric and traditional distance measures such as walking distance and straight-line distance.

This chapter was originally published in Environment and Behavior[1]. As a co-first author of this paper, I proposed and developed the idea of zone overlap, and theorized the idea with Festinger's functional distance. I also implemented the software to calculate and visualize the new metric, and calculated all spatial variables used in this study.

Chapter 4 introduces a novel approach for subspace decomposition, which is an essential process in many building layout analyses as well as one of the fundamental descriptors in morphological analysis. The two new metrics, zone overlap and sociospatial betweenness, proposed in Chapter 2 and Chapter 3 also require subspace decomposition. After we manually decomposed the layouts of the buildings in Chapter 2 and Chapter 3, we found that the study on subspace decomposition is relatively thin even though there is a critical need for the process and manual application to a large built environment is prohibitively time consuming. The essay demonstrates the problems of existing methods and relates them to the context of general network theories. Then we propose a new

[1] Kabo, F.*, Hwang, Y.*, Levenstein, M., & Owen-Smith, J. 2013. "Shared paths to the lab: a sociospatial network analysis of collaboration." *Environment and Behavior*, 0013916513493909.

decomposition method based on Newman's modularity (2006) and illustrate how the previous problems are solved with the new method. This essay was originally presented at the 9th International Space Syntax Symposium 2013[2].

Finally, Chapter 5 summarizes the major contributions of the essays and how the contributions of this dissertation fill a gap in this field. The chapter also suggests implications for the design of workplaces. Lastly, it discusses several limitations of the current study and the direction of further research work.

[2] Hwang, Y. 2013. "Network Communities in the Visibility Graph: a New Method for the Discretization of Space." In *Proceedings of the 7th International Space Syntax Symposium*. Seoul.

CHAPTER 2

# Spatializing Social Ties:

## Sociospatial betweenness centrality, a predictor of interaction diversity in workplace[†]

## 1. Introduction

A considerable amount of literature has been published on how space is related to the dynamics and outcomes of face-to-face interactions. These studies can be classified into two categories. The first category investigates how distance between two people is associated with the formation of a social tie between them (Allen and Fustfeld 1975; Festinger, Schachter, and Back 1950; Kahn and McGaughey 1977; Kraut, Egido, and Galegher 1988; Lee et al. 2010; Priest and Sawyer 1967). The second category of study investigates how one's advantageous or disadvantageous location in a building affects one's communication patterns. In most of such studies, a space is considered as an advantageous location when the space has small average distance to other spaces

[†] This research was supported by funding from the National Science Foundation (Grant No. 0724675).

although the definition of distance varies: the number of turns (Wineman, Kabo, and Davis 2008; Penn, Desyllas, and Vaughan 1999), the number of spaces (Toker and Gray 2008), and walking distance (Sailer and McCulloh 2012).

So far there has been little discussion about methods for quantifying space or evaluating a location's spatial advantage for space-communication studies. As we see, most of studies use distance or average distance to operationalize space's effect. Our approach proposed in this paper evaluates a space's communicational advantage not based on distance, but how much people's movement is concentrated on the space. This makes us to consider not only a space's location in the spatial network but also people's locations and how they are connected in their social network. To consider both of structures, we integrate the spatial network of a building layout and the social network of building occupants by spatializing social ties and then calculate socio-spatial centralities of a space from the integrated network.

We tested the validity of our novel approach with an empirical dataset of interaction patterns in a workplace collected using location tracking sensors. Our new sociospatial centrality measure is examined in association with a space's potential for attracting diverse people to engage in face-to-face communication, which is believed to be an essential ingredient of organizational innovation (Burt 2004; Obstfeld 2005). This measure, then, is compared to choice, a traditional betweenness centrality measure for spatial networks.

# 2. Background

## 2.1. Integrating Space and Social Network Analysis

### 2.1.1. Problems of using only spatial networks

Studies of the effect of physical environment on face-to-face communication often argue that higher integration (shorter average distance to other space) predicts more communication (Penn, Desyllas, and Vaughan 1999). This may be a plausible argument to some degree. As the famous Allen curve showing decrease of communication between employees as the distance between them increases (Allen and Fustfeld 1975), we may guess that a person with lower distance to all other people would have more chance of communication on average. Then, would the average distance to all other space be a good predictor for the average distance to all other people? The answer would be yes if people are distributed evenly on a floor. In such a floor, the average distance to all other space may show a good association with communication frequency as it works as a proxy to the average distance to all other people. If this is the case, it may be better to use the average distance to all other people instead of its proxy, unless we do not have enough information on people's locations.

Some may argue that a space with high integration, or low average distance to other spaces, accommodates more communication not because the space is closer to other people, but because such a space tends to be occupied by a person with higher social network power such as many social ties. This may be true. To verify this argument, we need to investigate the social structure among the people in the floor and how the social

structure is overlaid on the spatial structure. In other words, we need an integrated approach for the spatial network analysis and the social network analysis.

## 2.2. Traditional approaches for integrating the two networks

### *Spatializing nodes*

There have been attempts to add spatial dimension to social network analysis. One strand of such attempts is to investigate the effect of distance on the formation of a social tie (Festinger, Schachter, and Back 1950; Allen and Fustfeld 1975; Butts et al. 2012; Daraganova et al. 2012). Another approach is to use distance between two people as a weight of the social tie between each pair in the social network (Brandes 2008; Hipp, Faris, and Boessen 2012). In both cases, nodes are spatialized in the physical space and their metric distance between nodes is measured and assigned to the ties.

Both of the two approaches shrink spatial structure to the set of distances, overlooking how distances are organized. However, delicate differences in the socio-spatial situation related to the dynamics of communication behavior are difficult to capture only with distance. Let's think about two imaginary layouts accommodating the same social structure composed of three people A, B, and C (Figure 2.1). The three people occupy the same location across the two settings. And all pairwise distances A-B, B-C, and C-A have the same distance while the spatial structures are different.

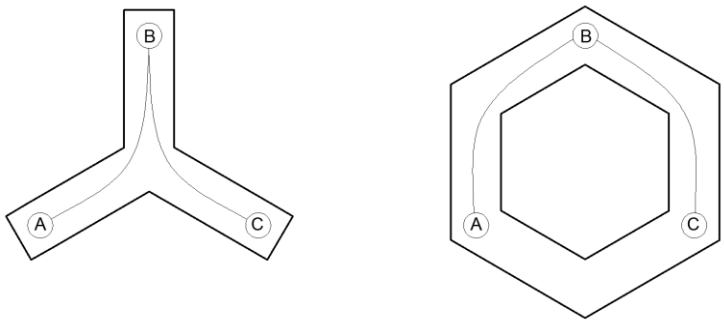



**Figure 2.1.** Limitation of Pairwise Distance. The three people occupy the same location across the two settings. And all pairwise distances A-B, B-C, and C-A have the same distance while the spatial structures are different.

Although pairwise distances are all the same, we may expect a different effect on social tie formation among the three people because of the different spatial structures around them. If all other things are equal across two settings, A and C is more likely to be exposed to each other in the left panel than the right panel. This is because the path from A to B and the path from B to C considerably overlap in the left panel while they are separated in the right panel. So A (or C) is more likely to see C (or A) when he or she goes to B. Such a difference in the two settings cannot be captured by measuring distances. The key is how movements of people are organized by the spatial structure.

***Integrating the two networks in a statistical model***

Another approach for investigating the joint effects of a social network and a spatial network is to use social centralities and spatial centralities as regressors in a regression model (Peponis et al. 2007; Wineman, Kabo, and Davis 2008). For example, to investigate the effect of social location and spatial location on one's work performance, a researcher may set up a regression model with independent variables including social centralities and spatial centralities that are separately calculated from each network.

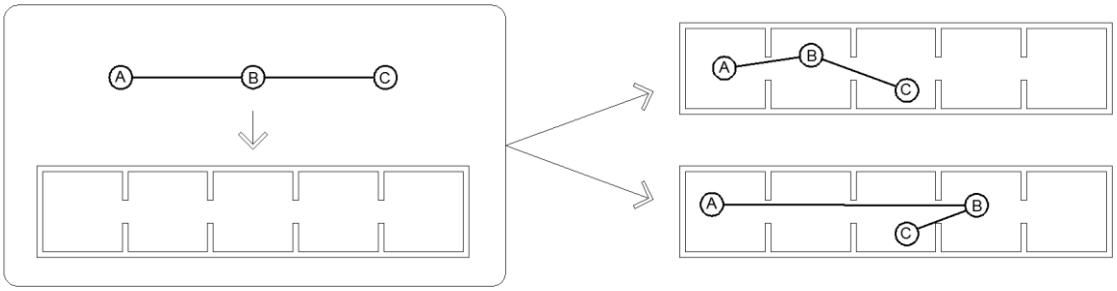



**Figure 2.2.** Same Spatial Structure, Same Social Structure, Different Situation. The same social structure is mapped onto the same spatial structure in a different way.

This approach, however, sometimes misses the delicate joint effect of the two networks. Figure 2.2 shows two different socio-spatial situations from the same social structure and the same spatial structure. But the way the social structure is mapped onto the spatial structure is slightly different across the two situations. Only B's spatial location is changed while A and C keep the same locations. Although B's spatial location is changed, B's spatial centralities are the same across two situations because the rooms occupied by B in the upper panel and in the lower panel are symmetrical in the spatial network. Thus, the spatial centralities of A, B, and C are the same across the two situations. The social centralities of three actors are also the same across two panels because there is no change in the social network. So a statistical model with both centralities will be exactly the same across two panels. It means that synthesizing the two networks' centralities in a regression model cannot distinguish the two different socio-spatial situations.

The socio-spatial situations of the two panels are quite different especially when we focus on face-to-face communication behaviors. For example, other things being equal, we may

expect a social tie between A and C would be more likely to form in the lower panel's socio-spatial situation than in the upper panel's situation even though the distance between A and C is the same across the two situations. This arises because the lower panel's situation would provide more opportunities for A to be exposed to C.

## 2.3. Spatializing Social Ties

To accurately capture the joint effect of a social network and a spatial network, we need a method to describe the combined structure of the both networks. This study proposes a method spatializing a social network on a spatial network. First, we overlay the social network of building users on the spatial network of the building. Then, actors in the social network are placed on their primary locations in the spatial network. Second, a tie connecting two actors' primary locations is placed following the path[3] in the spatial network connecting the two primary locations. This process is repeated until all ties in the social network are placed and we have the complete spatial projection of the social network (SPSN) of the floor (Figure 2.3). The key idea of SPSN is that it spatializes not only social actors but also social ties among them following the spatial structure.

SPSN provides a unique perspective different from the perspective a spatial network gives in terms of socio-spatial situations. The floor in Figure 2.3 is symmetric both horizontally and vertically. From the perspective of the spatial network, there is no difference in the upper part and the lower part, and in the left part and the right part. In

[3] We use the shortest spatial path for connecting two actors. Although people do not always follow the shortest path, people prefer shortest paths in an indoor space (Sailer and McCulloh 2012). We may use a complicated model for path choice respecting multiple possible paths, but we start with a simpler model using the shortest spatial path.

contrast, SPSN recognizes the differences because more people are located in the lower part. Also, SPSN recognizes the difference in the left and the right as seen in Figure 2.3 although the numbers of people in the left part and the right part are the same. Such a difference comes from B who has the highest social network power and is located in the left part.

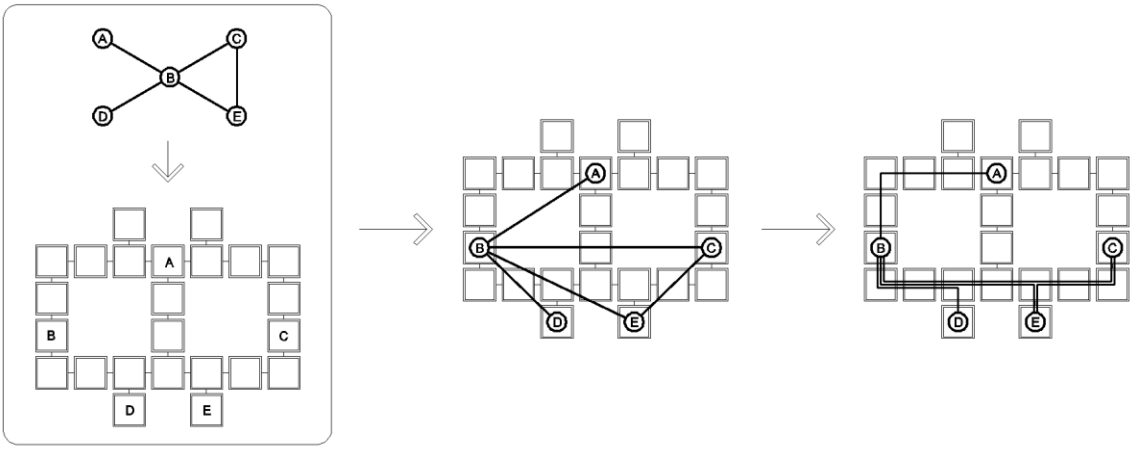



**Figure 2.3.** Spatial Projection of the Social Network (SPSN). It spatializes not only social actors but also social ties among them following the spatial structure.

SPSN provides a unique perspective different from the perspective a spatial network gives in terms of socio-spatial situations. The floor in Figure 2.3 is symmetric both horizontally and vertically. From the perspective of the spatial network, there is no difference in the upper part and the lower part, and in the left part and the right part. In contrast, SPSN recognizes the differences because more people are located in the lower part. Also, SPSN recognizes the difference in the left and the right as seen in Figure 2.3 although the numbers of people in the left part and the right part are the same. Such a

difference comes from B who has the highest social network power and is located in the left part.

Another unique advantage of SPSN is that it reveals how social ties between people overlap in the floor. The analysis of the spatial network would choose the central corridor as the space for most encounters because the corridor has the highest centrality. However, SPSN suggests a different picture. SPSN would choose the corridor in the bottom left where three different ties are overlapped and hence many opportunities of encounters exist. This observation lead us to the concept of sociospatial centralities, how centrally a space is located with consideration of a spatial network as well as a social network.

### 2.3.1. Sociospatial betweenness

SPSN enables us to formalize sociospatial centralities in a similar manner as the general network centralities are defined. For example, closeness centrality of a person in a social network measures how easily a person can reach other people and usually is defined as the inverse of the average social distance (number of ‘hops’ in the network) from the person to all other people in the network (Jackson 2008). In a similar way, we can define sociospatial closeness centrality (SSC) of a space measuring how easily people can reach the space as the inverse of the average walking distance from the space to all other people. Also, degree centrality in a social network is defined by the number of socially connected people (neighbors) to the actor, measuring how well connected an actor is and (Jackson 2008). Similarly, we may define sociospatial degree centrality (SSD) as a measure of how sociospatially well-connected a space is by counting the number of spatial neighbors, people within a certain spatial distance from a space of interest.

Among many possible sociospatial network centralities, we focus on sociospatial betweenness centrality (SSB) in this study. Betweenness centrality is a metric for how well located a node is in terms of connecting to other nodes (Jackson 2008) and usually is defined as the fraction of shortest paths that pass through the node out of all possible shortest paths between node pairs (Newman 2005). Expanding this concept, we propose sociospatial betweenness centrality of a space as a metric for how well situated a space is in terms of connecting other people. It is defined by the fraction of social ties passing through a space of interest in the SPSN of a given building layout (Figure 2.4).

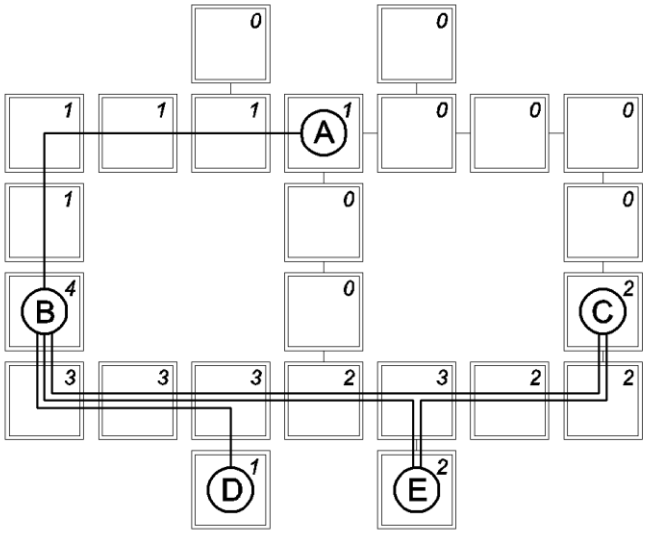



**Figure 2.4.** Sociospatial Betweenness (unnormalized). Numbers at the upper-right corner of each space denote the number of social ties passing through the space.

Betweenness centrality is a popular measure in network analysis of many fields. Compared to closeness centrality that has been intensively used for spatial network analysis in the name of ‘integration’, betweenness centrality has been rarely used for

analyzing the spatial network of indoor spaces except for a few studies (Kembel et al. 2014; Li, Claramunt, and Ray 2010).

One of the reasons for the rare use of betweenness centrality in the analysis of indoor space is that the value of betweenness centrality is significantly influenced by how corridor spaces are decomposed. Let's think about a layout where a corridor connects two offices occupied by two people, P and Q, respectively. As seen in panel (A) of Figure 2.5, the betweenness centralities of corridor spaces change as the corridor is decomposed into smaller spaces. As the corridor is decomposed into three subspaces from one space, the central corridor space's betweenness centrality becomes 16% higher than those of the corridor spaces at both ends. When the corridor is decomposed into 4 subspaces, the gap becomes almost 40%.

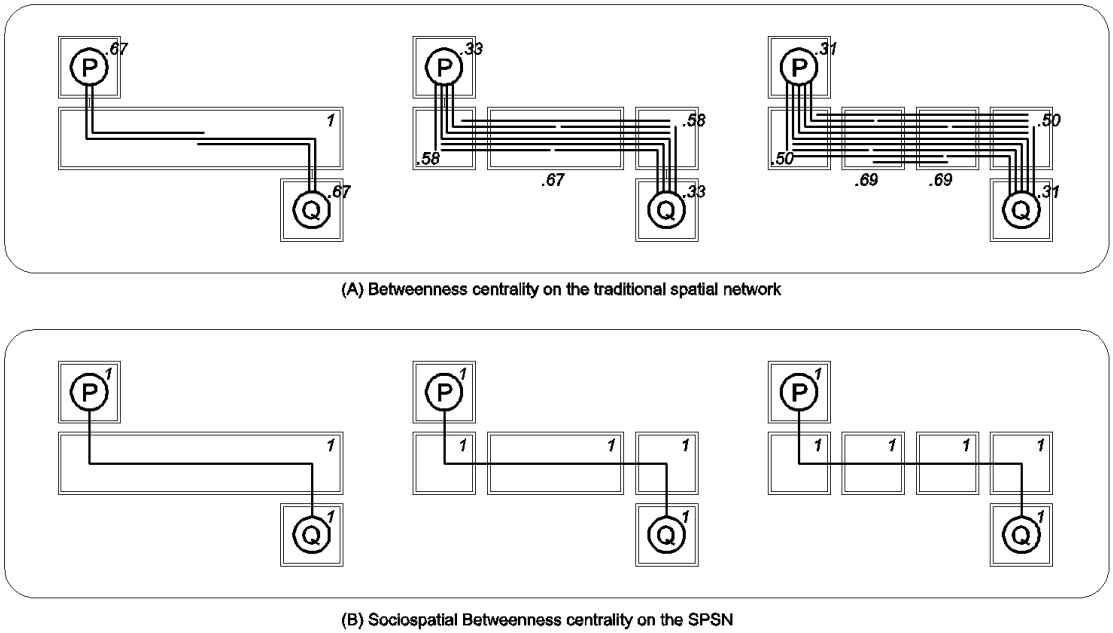



**Figure 2.5.** Problem of Betweenness for Indoor Space. The betweenness centralities of corridor spaces on the traditional spatial network change as the corridor is decomposed into smaller spaces (Panel A).

Such unexpected noise happens because betweenness centrality in a spatial network is determined by the number of paths, or traffics from all spaces to all other spaces. So as the corridor is decomposed into more subspaces, it will generate more ‘internal traffic’ within the corridor that has a tendency to inflate a central zone’s betweenness. In contrast, sociospatial betweenness is robust to the way a corridor is decomposed. As seen in panel (B), sociospatial betweenness does not change no matter how the corridor is decomposed. This is because ‘traffic’ is generated by only social actors, not by spaces in SPSN.

# 3. Method

## 3.1. Hypothesis

Our central claim is that the dynamics of face-to-face interactions in workplace are better understood when the workplace’s spatial structure and the social structure are considered together. As one of integrative tools for analyzing sociospatial structure, we propose sociospatial betweenness, the fraction of spatialized social ties passing through a focal space. If people visit their social neighbors more often than unrelated people, a space with high sociospatial betweenness would be traversed by many different people as compared to a space with low sociospatial betweenness. Then, we expect that a space with high sociospatial betweenness would encourage encounters and interactions among diverse people instead of frequent encounters and interactions among a handful of people. We also expect that sociospatial betweenness of a space is more significantly associated

with the diversity of interaction of the space than betweenness of the space calculated from the spatial network. Thus,

**Hypothesis 1 (H1):** The greater the sociospatial betweenness of a space, the more diverse the interaction that happens in the space.

**Hypothesis 2 (H2):** The effect of spatial structure on interaction diversity would be associated more significantly with sociospatial betweenness than spatial betweenness.

## 3.2. Dataset

***Sample Selection.*** The study group is a department of an international manufacturing company. The group was comprised of 36 professionals occupying the whole floor of a multi-story office building located on a larger campus of the corporation's buildings. Managers are occupying closed perimeter offices, non-managers are assigned to partitioned offices.

***Interaction.*** To test the hypotheses, we used a dataset indicating the locations of interactions among the group. All people in the floor participated in data collection during working hours for the 9 week study period in 2009. The dataset collected interactions through an indoor positioning system using ultra-wideband (UWB) location tracking system. From the signals indicating each person's moving trajectory, we extracted interactions among the professionals following the definition: an 'interaction' is an event when the distance between two people (tags) was maintained within 10 feet walking distance for more than 10 seconds' duration. Each interaction in the dataset has the location, timestamp, and duration of the interaction with personal IDs participating

the interaction. The dataset's validity was cross-checked with the sociometric survey and on-location interviews. For further details, see Wineman et al. (2014).

***Construction of spatial network.*** We constructed the spatial network of the floor where we collected the interaction dataset by decomposing the floor into subspaces and connecting adjacent spaces as follows. First, primary assigned spaces such as offices and public spaces such as break rooms, meeting rooms, and restrooms were treated as discrete elements. Second, connector spaces such as hallways were decomposed to identify paths between professionals' primary spaces. To achieve this goal, the connector spaces immediately adjoining the doors and openings to offices were demarcated as thresholds. Then, connector spaces between thresholds were subdivided into smaller spaces so that the distances between the centroids of the resulting spatial element reflected actual walking distances, conditional on the arcs or edges connecting these centroids not crossing walls or other physical barriers. We identified total 221 spaces including office spaces, common spaces, and corridor spaces.

***Spatializing social network.*** Nodes representing 36 professionals were placed in their primary assigned spaces. From a node, its social ties were placed following the shortest spatial paths to its connected nodes. This process was repeated for all 36 professionals. In this study, we used the complete graph for the organization's social network because all professionals participating in the study belonged to the same unit in the same floor.

***Spaces of study***. Out of total 221 spaces, we included only corridor spaces (131 spaces) in our dataset. We have several reasons for this. First, the focus of the study was on unplanned encounter and interactions in corridor spaces tend to be more unplanned than interactions in office spaces. Furthermore, unplanned interaction is more affected by

spatial structure (Allen 1977). Thus, corridor spaces are better places to observe the effect of spatial structure. Second, all end nodes have the same value for betweenness centrality by definition. All office spaces in this study are end nodes because an office space is connected to corridors with only one link. Thus betweenness centrality of all office spaces would show no difference. Third, the signals for the location tracking system could not reach some of the offices in the wing spaces. Because of these reasons, we excluded non-corridor spaces from our analysis. Excluding non-corridor spaces does not mean that we created the spatial network of the floor only with corridor spaces. The spatial network is constructed with all spaces, and variables are calculated with all spaces. Then only corridor spaces are included in the dataset for the statistical analysis.

## 3.3. Variables

Our dependent variable is ***Interaction Diversity***. For each corridor space, we counted the number of people who engaged in communication in the space over the study period. For example, if two people had a conversation in a space 10 times over the study period, interaction diversity of the space is two; if 10 people had conversation in a space only one time over the study period, interaction diversity is 10.

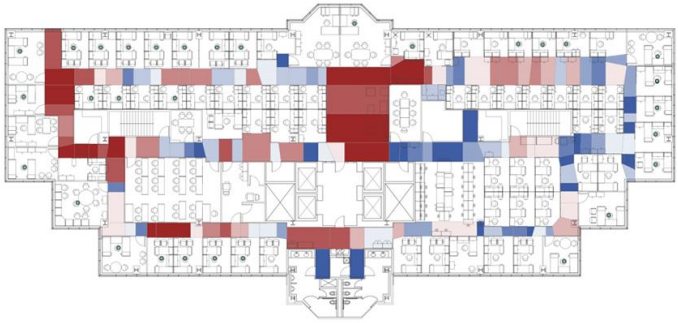
We have two independent variables: ***Sociospatial Betweenness (SSB)*** and ***Spatial Betweenness (SPB)***. For each corridor space, we calculated the fraction of social ties passing through the space among all social paths between pairs, which we call SSB. SPB of a space is the betweenness centrality of the space in the spatial network of the floor.

**Table 2.1.** Correlations among the variables

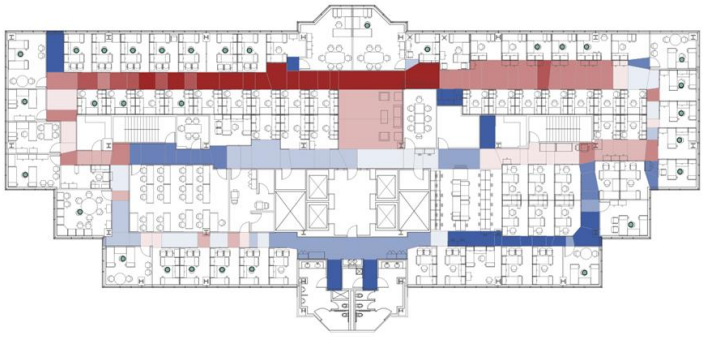
| | Variables (N=131) | Mean | SD | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|
| 1 | Interaction Diversity | 10.870 | 7.430 | | | | |
| 2 | Sociospatial Betweenness | 0.319 | 0.144 | 0.324 ** | | | |
| 3 | Spatial Betweenness | 0.129 | 0.067 | 0.151 | 0.514 ** | | |
| 4 | Coffee Distance | 32.798 | 17.392 | -0.434 ** | -0.106 | -0.102 | |
| 5 | Area | 3.037 | 2.478 | 0.423 ** | 0.111 | 0.334 ** | -0.046 |

*$p < .05$. **$p < .01$. ***$p < .001$

Our previous study showed that coffee bar is a strong social space (Wineman et al. 2014) and the coffee bar tends to draw people from across the floor. To control for the effect of coffee bar, we created a variable, ***Coffee Distance*** capturing walking distance from a space to the coffee bar in meters. Another control variable is ***Area***, the size of a space in square meters because a large space tends to have more interactions in the space.



(A) Interaction Diversity



(B) Sociospatial Betweenness

(C) Spatial Betweenness

**Figure 2.6.** Distributions of the Key Variables (red: high, blue: low). Circular dots denote the location of professionals.

# 4. Findings

Negative binomial regression were applied to examine the contributions of locational advantage as measured by the variable Sociospatial Betweenness (SSB) and Spatial Betweenness (SPB) to our dependent variable, Interaction Diversity. We constructed three regression models. Model 1 was run with SSB and the control variables (Coffee Distance and Area). Model 2 was run with SPB and the control variables. Finally, Model 3 was run with SSB, SPB, and the control variables.

SSB is positively and significantly related to Interaction Diversity both in Model 1 and Model 3, confirming H1. A space with higher SSB tends to have more diverse interaction partners as we expected. Locational advantage of a space in a sociospatial structure increases the diversity of interactions that occurred in the space.

In contrast, SPB shows no significance in Model 2, confirming H2. Although SPB shows significance with SSB in Model 3, it is negatively related to Interaction Diversity against our expectation. As we see in Figure 2.6, SPB is high in the central zone and gradually decreases as a space becomes peripheral. However, Interaction Diversity does not show such a pattern. Instead, corridors in the peripheral zones show high diversity when they are surrounded by many people or adjacent to key functional areas such as the coffee bar. SPB is not a good measure to capture how dense occupation is surrounding a space because it treats all spaces equally whether a space is occupied by a person or not. In contrast, SSB treats only occupied spaces as sources and destinations, and hence it is sensitive to the locations of people.

**Table 2.2**. Effects of Betweenness Centralities on Interaction Diversity

| DV=Interaction Diversity | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Sociospatial Betweenness (SSB) | 0.0241*** | | 0.0329*** |
| | (0.0000) | | (0.0000) |
| Spatial Betweenness (SPB) | | -0.0026 | -0.0192* |
| | | (0.7170) | (0.0167) |
| Coffee Distance | -0.0365*** | -0.0389*** | -0.0367*** |
| | (0.0000) | (0.0000) | (0.0000) |
| Area | 0.0290*** | 0.0332*** | 0.0347*** |
| | (0.0000) | (0.0000) | (0.0000) |
| Constant | 2.1478*** | 2.5921*** | 2.2199*** |
| | (0.0000) | (0.0000) | (0.0000) |
| Observations | 131 | 131 | 131 |
| 2xLog-likelihood | -803.93 | -815.31 | -798.43 |
| AIC | 813.93 | 825.31 | 810.43 |

Note. Standard errors in parentheses.
*$p < .05$. **$p < .01$. ***$p < .001$
DV=Dependent Variable

The difference between SSB and SPB is clear in the comparison of the top right corridor and the middle right corridor in Figure 2.6. SSB is much higher in the top corridor than in the middle right corridor because more people are adjacent to the top corridor and hence more people pass through it. But there is no big difference in SPB between the two corridors because there is no big difference in the number of spaces attached to each corridor and the paths passing through it. By comparing these two corridors we can see how SSB reflects actual communication patterns more accurately; the top right corridor shows higher Interaction Diversity than the middle right corridor as seen in Figure 2.6.

The control variables performed as we expected. Coffee Distance shows consistent and strong significance across models as in the previous study (Wineman et al. 2014). Coffee

bar is a location that brings diverse people together into conversations. Also, a larger space tends to accommodate more interaction partners largely because of its size.

# 5. Conclusion

One of the more significant findings to emerge from this study is that the effect of space on face-to-face communication is better understood with an integrative approach to the social structure and spatial structure. For such an integrative approach, we proposed a novel method, the spatial projection of a social network (SPSN) obtained by spatializing social ties in the social network as well as social actors. From the projection, we derived Sociospatial Betweenness and have argued that Sociospatial Betweenness is a better instrument to explain a space's effect on communication patterns, such as interaction diversity, compared to the counterpart centrality of a spatial network. From the SPSN, we tested only Sociospatial Betweenness in this study. It would be interesting to explore the possibilities of other sociospatial centralities derived from the SPSN such as sociospatial degree and sociospatial closeness.

Another contribution of this study is that this study reinvigorates betweenness centrality that has long been underused in the analysis of building layouts in spite of its powerful concept and popularity in other fields. Sociospatial betweenness would give researchers in this field wider options of analyzing building layouts. We believe that it substantially widens the options because sociospatial betweenness is not based on distance and hence less likely to give redundant information when it is used with other distance based measures.

Our analysis is restricted to only one building; so it is problematic to make broad generalizations. However, in the previous study, an elementary form of sociospatial betweenness has been used to analyze the relationship between spatial structure and innovative outcome across three different layouts and it was found to have a significant effect in all three layouts (Wineman et al. 2014).

This research has raised many questions in need of further investigation. First, we used the shortest paths to calculate sociospatial betweenness of a space. This implicitly assumes that people always move following the shortest path in the spatial network when they visit others' offices. This might be too strong an assumption. We may relax the assumption so that shorter paths are more likely to be chosen while longer paths are less likely to be chosen as suggested in random walk betweenness (Newman 2005). Second, the SPSN reflects the overlap of movement induced by peoples' primary locations such as office spaces. In many cases, however, peoples' offices are not the only attractors that generate movement. The coffee bar, an elevator hall, a restroom, or copy machines can also be key attractors. We may extend our framework so that it can include such key facilities.

CHAPTER 3

# Shared Paths to the Lab:

## A sociospatial network analysis of collaboration[†]

Felichism Kabo*, Yongha Hwang*, Margaret Levenstein, Jason Owen-Smith.

## 1. Introduction

Social relationships shape the activities of organizations, teams and individuals in complicated ways (Burt 2004; Hansen 1999; W. W. Powell, Koput, and Smith-Doerr 1996), but social scientists are only beginning to explore systematically how the arrangement of physical space influences workplace interactions and outcomes. People work and interact in the built environment (Grannis 2009; Hua et al. 2010; Heerwagen et al. 2004). Research that disregards space or analytically divorces social phenomena from

---

[†] This paper is originally published in *Environment and Behavior* (2013). Kabo, F.*, Hwang, Y.*, Levenstein, M., & Owen-Smith, J. 2013. "Shared paths to the lab: a sociospatial network analysis of collaboration." *Environment and Behavior*, 0013916513493909.

* The first two authors contributed equally to the preparation of this article (from the author's note in the original paper published in Environment and Behavior).

location is likely to result in impoverished theories, biased findings, and misspecified models (Kono et al. 1998).

Many of the processes and outcomes that are the focus of organizational analysis depend upon social networks. The most elemental level of analysis for understanding networks based on information sharing, collaboration, or teamwork may be the dyad (Mizruchi and Marquis 2006). Despite the importance of dyads, little work has explicitly examined spatial effects on dyad formation (Allen and Fustfeld 1975; Sailer et al. 2009; Wineman, Kabo, and Davis 2008). Spatial effects have not been robustly incorporated in such social science models because we have lacked spatial measures that are nuanced enough to operationalize key concepts. This paper takes initial steps toward more fully integrating spatial and social explanations of collaborative relationships at work by developing a new measure of physical proximity that more effectively captures classic concepts of the effects of space on the likelihood of interaction (Festinger, Schachter, and Back 1950). We demonstrate that our measure, zone overlap or the extent to which a pair of individuals share common physical spaces, explains rates of collaboration formation among interdisciplinary life scientists working in two research buildings on the campus of a large public research university. The explanatory power of this measure is distinct from the effects of more traditional distance variables including the metric measures of straight line and walking distances and the topological measure of turn distance.

Most contemporary efforts to understand spatial effects in organizational settings employ physical distance as a proxy for the subtle ways in which proximity enables or hinders interaction (Cowgill, Wolfers, and Zitzewitz 2009; Liu 2010; Sailer and McCulloh 2012). While this body of research has incorporated spatial effects in organizational analyses,

their measure of distance cannot generally capture the powerful but subtle relational and topological effects of space that we refer to as functional proximity. We draw on classic work examining spatial influences on interaction (Festinger et al., 1950) and pioneering efforts to capture the relational aspects of the built environment (Hillier and Hanson 1984) to propose a new conceptualization of space, the functional zone, which captures individual spheres of operation in the workplace. From the functional zone concept we develop measures of zone overlap. Path and areal zone overlap*s* between individuals capture key aspects of space that increase or decrease the likelihood of dyadic interaction. Zone overlap measures offer continuous, quantitative indices of proximity that are robust across spatial layouts and thus offer the possibility of application and generalization across multiple organizational settings.

Space is the platform on which face-to-face social interactions and the networks that result from them are enacted. Nevertheless, efforts to develop systematic sociospatial organizational research have languished since seminal, but largely descriptive analyses (Allen 1977; Festinger, Schachter, and Back 1950). Festinger, Schachter and Back's study of interactions among residents in a new campus community for World War II veterans returning to university under the GI bill offer particularly valuable insights that have been little developed. This study drew a distinction between two critical mechanisms through which space shapes interaction. The first is *physical distance* which captures the costs (in time and effort) of interaction for a particular dyad. Here, the assumption is that greater distances between people make it more difficult to initiate and sustain face to face interactions.

The second mechanism, which was dubbed *functional distance*, focused more explicitly on the relational aspect of physical layouts by emphasizing, for instance, the ease and difficulty of movement among spaces. The implications of functional distance for social and organizational research have eluded careful consideration and measurement. This paper operationalizes functional distance in terms of overlapping zones of activity and then compares those measures to metric and topological characterizations of physical distance.

We test the assertion that the arrangement of physical space exerts significant effects on collaboration and out starting point is the assumption that spatial effects are probabilistic and contingent rather than deterministic and universal (Sack 1986). For example, someone whose workspace is located next door to a popular coffee bar, favorite break space or even much visited rest room (Pfeffer 1992) might forgo the increased opportunities for interaction offered by her location through the simple expedient of shutting a door or wearing large noise cancelling headphones. Proximity need not beget interaction. Likewise the actual impact or importance of the costs imposed by physical distance may vary with the overall topology of the building where interactions happen. This suggests that the effects of physical distance will vary with building design, while functional distance will exert more consistent effects.

### *Collaboration and Space*

Festinger et al noted that brief passive or unscripted contacts constitute the foundation for new tie formation. A determinant of these chance encounters is what they referred to as

required paths, such as the one that an individual must take from home to the bus stop.[4] Potential dyad members are more likely to initiate contact to the extent that their required paths cross or overlap. Yet the absolute physical distance, say between homes, is a poor predictor of potential path overlap. Path overlap is better predicted by the relative positions of the multiple spaces in which people routinely navigate. In other words, if the overall configuration of a physical space and the distribution of commonly visited locations within it require individuals to encounter one another more often in the course of their daily activities, they will be more likely to interact, share information, and develop collaborative relationships.

The most notable contemporary method for configurational, system level analysis of buildings is *space syntax*. Space syntax techniques highlight the relational nature of space by converting physical layouts into networks that represent proximities among rooms and passageways in relational terms (Hillier 1996; Hillier and Hanson 1984). Like social networks, spatial networks can be used at multiple levels of analysis, for example, buildings, campuses, and cities. At the level of buildings, spatial networks closely mirror their social counterparts as they allow for egocentric, dyadic, and overall network levels of analysis in a specific spatial system (see online Appendix Table A1). This paper seeks to expand our understanding of the sociospatial dynamics of collaboration networks at the dyadic level. But it is precisely at the dyadic level of analysis that space syntax's contributions to the development of a sociospatial science start to diminish.

[4] In this example, the path is only "required" if the individuals in the dyad both take the bus and not alternate forms of transportation such as cars or bicycles. Whether one follows the presumed path depends on social, economic, and cultural factors. But given all of those, the overlap of paths affects the probability of interaction.

Dyadic topological distance measures derived from space syntax are more likely to be highly correlated with metric distance at the micro level of buildings as opposed to the more macro level of cities and regions. Thus, we make independent, dual comparisons of our zone overlap measure with metric and topological physical distance measures with respect to capturing the potential for dyadic and unplanned face-to-face encounters. That is, we test the proposition that zone overlap better captures the effects of space on collaboration dynamics than do physical distance measures.

### *Measuring Functional Proximity*

Following Festinger et al, efforts to examine the impact of space in organizational processes were rather coarse-grained. For reasons beyond the scope of this study, practical conceptualization of functional distance or relational aspects of space has lagged the use of physical distance, even though functional distance is arguably better at capturing the latent interactions between actors in a specific spatial environment. The incorporation of space in these studies has been mostly limited to physical distance, especially the simpler straight line or "as the crow flies" distance even relative to the more nuanced measure of walking distance (Monge and Kirste 1980). A simple example highlights differences in straight line versus walking distances in the analysis of physical spaces. In Figure 3.1, where each arc has a unit length, the actual or walking distance between individuals A and B is two units while the straight line distance is 1.414 units. The disparity between walking (five units) and straight line (one unit) distances is even greater for individuals A and C.

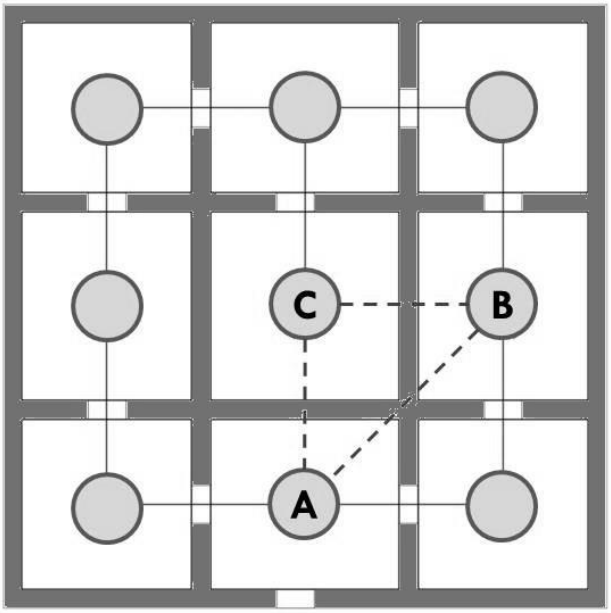



**Figure 3.1.** An illustration of straight line and walking distances between individuals A, B, and C in a simple spatial layout. For simplicity, individuals are restricted to orthogonal movements. Walking movement paths are depicted using solid lines while the straight line paths are shown as hatched lines.

Walking distances between the primary spaces (e.g. offices) that individuals occupy offer more salient conceptions of distance than do straight line measures. Nevertheless, point to point walking distances still miss aspects of space that shape the likelihood of passive contacts in the course of normal daily activity. Consider Figure 3.2, which presents several scenarios where building features might alter the likelihood that the occupants of two offices will encounter one another. Panel A shows the walking distance from door to door for the two offices. In subsequent panels the walking distance between offices remains constant, but the placement of stairwells alters the likelihood that occupants will encounter one another.[5] Panel B represents a configuration where stairwells at the ends of each corridor would lead office occupants to enter and depart by different paths, lowering

[5] For the purposes of this discussion, we assume that people typically enter and leave their offices by way of the nearest stairwell.

the chance that they encounter one another. Given the elbow shaped bend in the hallway it is possible that occupants might rarely even see one another. Panel C, in contrast, places a single stairwell equidistant from the two offices. Office occupants are likely to encounter each other at the stairwell, but soon part ways as they head to their separate spaces. Panel D suggests an even greater likelihood of passive contacts. Here occupants may meet at the stairwell and sometimes walk together briefly as their paths overlap. In addition, one person's path to the stairwell will lead them by the other's office door. In this configuration, then, the possibility of passive contact does not depend entirely on coordinated comings and goings via the stairwell. Finally, consider Panel E, which seems to us to offer the greatest possibility for passive contact. Panel E features a shared stairwell, a passed door, and a longer walking path overlap than in Panel D. In these alternative layouts, the walking distance between offices remains constant but their occupants' functional zones vary dramatically in ways that introduce greater or lesser possibilities for unplanned, face-to-face encounters in the course of daily interaction. Walking distance is a richer conception of physical distance than straight line distance, but is limited in its ability to capture how these more subtle effects of spatial layout effect functional distance.

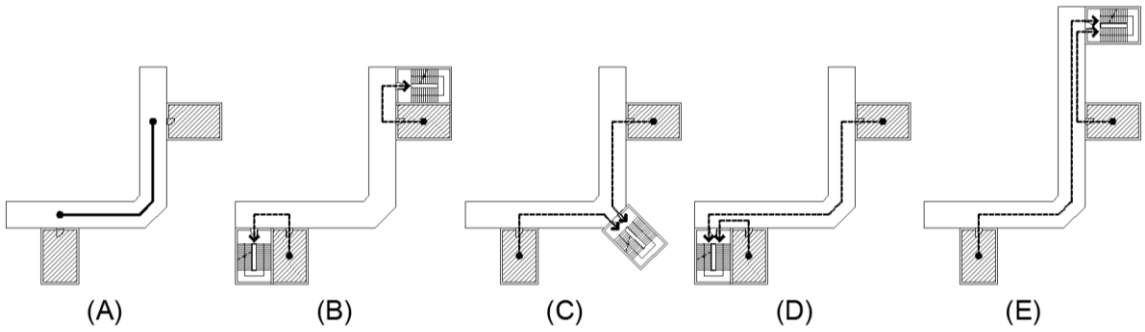



**Figure 3.2.** Further demonstration of the limitations of physical distance as a robust proxy for the finer-grained effects of spatial proximity. Holding the physical distance constant, changes in the relative locations of the individuals in the latent dyad lead to dramatic differences in the expected likelihood of encounters or interactions between them.

Festinger et al defined functional distance in terms of the "positional relationships and features of design" that make it more or less likely that two individuals will have unscripted encounters or interactions (Festinger et al., 1950, pp. 34-36).[6] This implies that the distance refers to topological relationships between spatial elements. To highlight this relational meaning, we substitute *proximity* for *distance.* Not only is proximity understood to be the antonym of distance ("Merriam-Webster's Collegiate Dictionary" 2003), but it also encompasses the broader dimensions of adjacency and contiguity. Therefore, from this point on, we will refer to *functional proximity* whenever we mean to invoke the *functional distance* of Festinger et al. In accord with previous usage of the term "functional proximity," our concept shares the connotation of accessibility between actors engendering interactions between people (Moodysson and Jonsson 2007; Pierce, Byrne, and Aguinis 1996; G. N. Powell and Foley 1998). However, our construct is

[6] In any spatial environment, individuals take certain paths to and from their primary spaces however these are defined. The emphasis here is on the likelihood of encounters between individuals given the paths they are likely to take in their specific environments. Individual, organizational, and sociocultural factors play a vital role in determining whether potential ties are consummated into actual relationships. That is an important question but is neither the focus of this study nor a precondition for the salience of the zone overlap concept. It will influence the relationship between zone overlap and collaboration in different social contexts, and this is an important topic for future study.

analytically more precise and quantifiable, lending itself to application in empirical and comparative studies. We recognize that the romantic relationships in the Mainiero, Pierce et al, and Powell and Foley studies are driven by different rationales than the research collaborations in our study. However, these studies explicitly apply the functional proximity concept while research on the phenomena we are interested in – workplace and scientific collaborations – does not (Heerwagen et al. 2004; Toker and Gray 2008).

To say that two individuals are proximate is to infer a degree of closeness between them on the basis of contiguity in a specific dimension. Individuals in the workplace have more or less established spheres of operation. An individual might always take the same elevator or stairway to their office, use one restroom over another, or prefer to take breaks in a specific area. We define the individual's sphere of operation as the functional zone. It is an aggregate function of the spaces that are the sites of task performance or personal movement in the workplace.

For the biomedical research buildings analyzed in this study, we emphasize four types of spaces: individuals' workspaces (offices, labs), public or shared spaces (restrooms), circulation spaces (elevators, stairways), and connectors (hallways). Of course one could draw up a different typology of spaces for these focal buildings, and it is likely that in buildings supporting other kinds of work, such as engineering production or software research, employees' functional zones will consist of sets of spaces that differ from the four types outlined above. In other words, our typology is not necessarily exhaustive or general enough to apply in its entirety across different building types or usages. Using the four types of spaces, we consider each individual's functional zone to be bound by their individual workspaces, most proximate restrooms, and the closest elevators, and threaded

together by the connector spaces. For simplicity, we also assume the individuals take the shortest path available. Our definition of the individual's functional zone is therefore quantifiable and provides metrics that allow for the capture of spatial use patterns at the individual level.

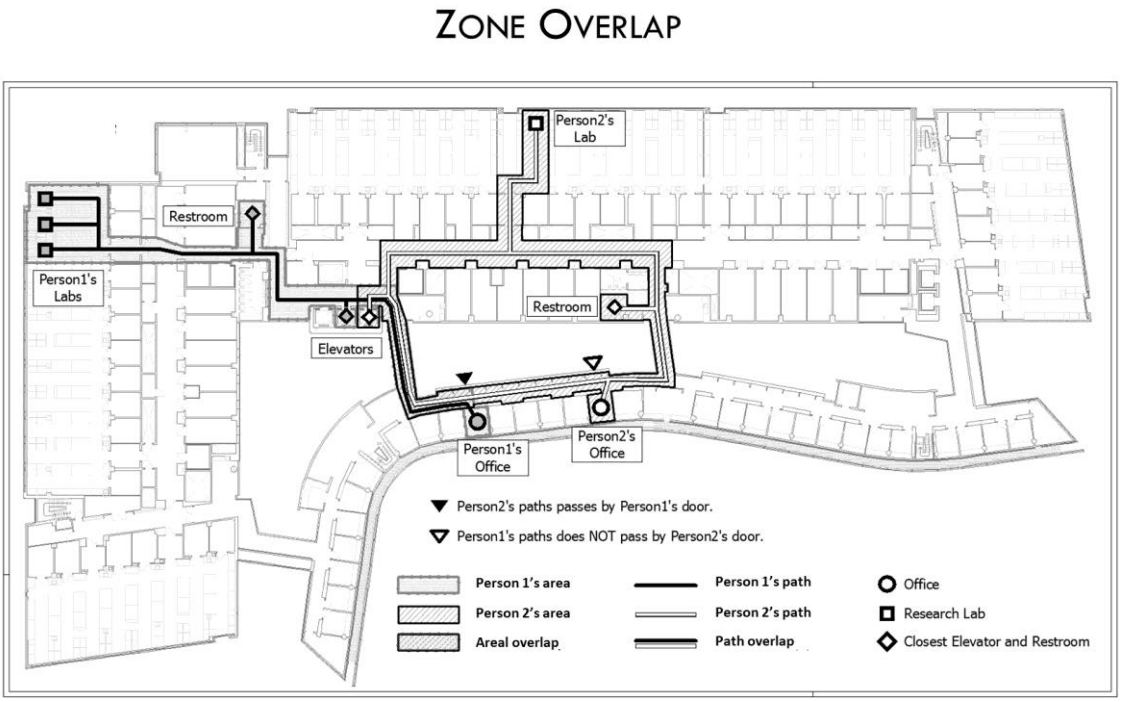



**Figure 3.3.** An illustration of the two measures of zone overlap (areal and path) using the BLD1 building. Also shown is the related concept of "door passing." The shared spaces that bound each person's functional zone in the example above are the elevators and the restrooms.

Consider Figure 3.3, which represents the work paths of two hypothetical investigators who share a floor in the BLD1 building. The path outlined by a heavy black line traverses the shortest walking routes connecting Person 1's assigned office, lab space, the nearest elevator and the nearest relevant restroom. The path depicted by the double gray line does the same for Person 2. While the offices assigned to these investigators (identified by circles) are very close together in both physical and functional terms, their spheres of

operation do not overlap to a great degree. Their overlap is represented by the path that includes both the black and double gray lines. These particular work paths overlap primarily because of a shared elevator, suggesting that these researchers are most likely to bump into one another when they are entering or leaving the building rather than during the course of their daily work or as they move back and forth between their offices and laboratory spaces. In this paper, we focus our attention simply on the extent to which paths overlap or not. Future work attending to the different roles of public and shared spaces (such as restrooms, break rooms, conference rooms, or scientific instruments) might offer even stronger insights into collaborative dynamics.

An individual's functional zone defines his or her sphere of potential interactions with others in a spatial system. It does not measure the impact of actions to constrain others' access to space. Functional zone should not be confused with territory as defined in the territoriality literature (Sack 1986; Sack 1993; Sykes 1977). Human territoriality represents a strategic intent to control or influence people and social interactions. For example, in the home parents might employ a territorial strategy by limiting children's access to a particular room. Similarly, zoning prescribes what activities are allowed within certain areas of a city (Sack, 1986). The crux of territoriality as a strategy is the intent to control differential access to material and human resources including social interactions. We use "zone" rather than "territory" in order to avoid confounding the impact of control over space with the probabilistic effects of simply being present in space.

Functional zone is an individual level measure that facilitates the development of dyadic and potentially group-level spatial measures that are not replications of physical distance.

The dyadic measure we propose is the zone overlap between individuals, which could be *path* or *areal* overlaps. *Path* measures of overlap correspond to the paths in individuals' functional zones while *areal* measures are contingent on the total size of the spaces in their functional zones. Whether path or areal, measures of zone overlap allow for dyadic and higher level analyses. This relational conception of proximity enables novel analyses of the dynamics and outcomes of interactions in sociospatial contexts.

## 2. Hypotheses

Our central claim is that space matters for the dynamics and outcomes of workplace interactions because proximity increases the likelihood of unplanned face-to-face contact while decreasing the costs of planned meetings. Physical distance by itself is a poor proxy for the role of space in social interactions and relations. Space also acts through adjacencies and contiguities, that is, functional proximity.

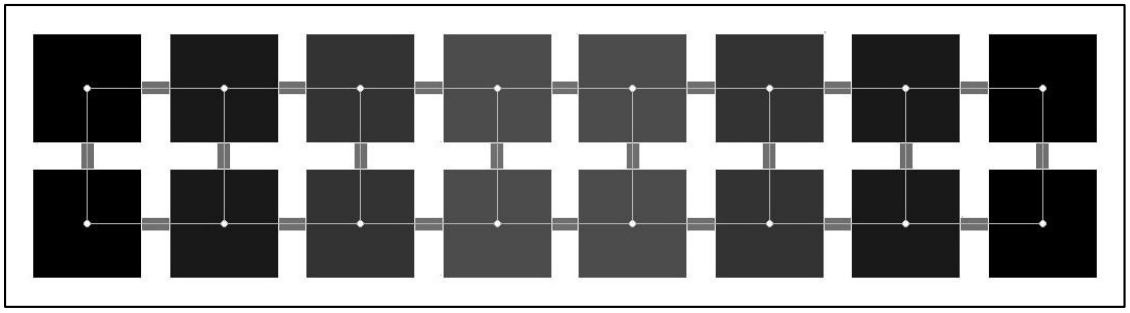


**Figure 3.4.** A more linear 16-space layout with the links between the spaces shown as light gray lines. From darkest to lightest, the spaces are coded according to their mean distance. The four values of mean distance are 4.267, 3.467, 2.933, and 2.667.

Physical distance affects the likelihood of interaction, but measures of physical distance are sensitive to topology or configuration effects. For example, consider Figures 3.4 and

3.5 which show two different 16-space layouts, the former linear and rectilinear and the latter square and compact, and where each space is a 4-unit square and the centroid-to-centroid distance is one unit in length. Calculation of the mean distance values shows that the spaces in Figure 3.4 (mean = 3.333, SD = 0.611) are generally at greater distances from each other than are the spaces in Figure 3.5 (mean = 2.667, SD = 0.377). This suggests that the physical distance between spaces is affected by the overall layout of the building or spatial system. Thus,

**H1a**: the greater the walking distance between two people the lower the potential for knowledge transfer between them and the lower their dyadic research collaboration index; this effect is more significant for linear layouts relative to more compact layouts.

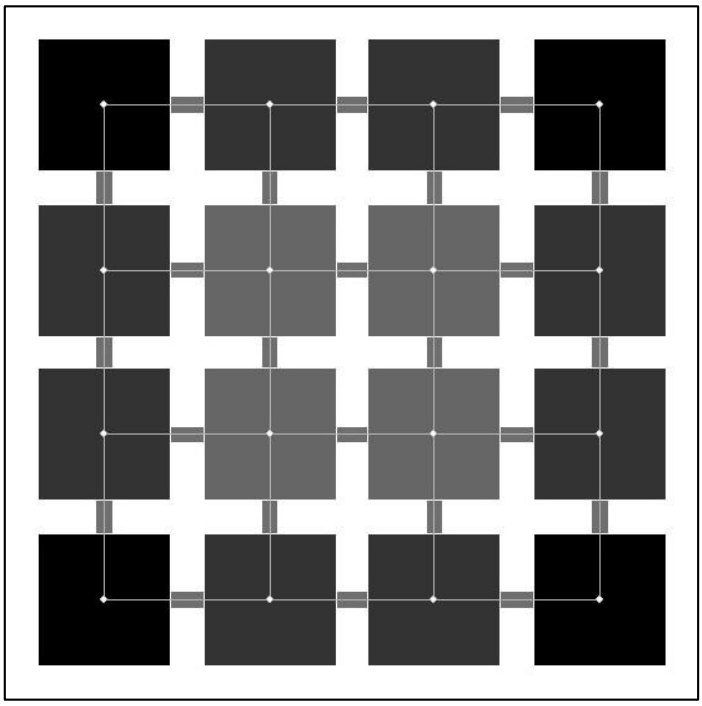


**Figure 3.5.** A more compact 16-space layout with the connections between the spaces shown as light gray lines. The spaces are coded from darkest to lightest to correspond with the three levels of mean distance (values are 3.200, 2.667, and 2.133).

**H1b**: the greater the turn distance between two people the lower the potential for knowledge transfer between them and the lower their dyadic research collaboration index; this effect is more significant for linear layouts relative to more compact layouts.

While functional proximity (zone overlap) is dependent to some degree on the physical distance between individuals, it emphasizes the relative locations and walking paths of individuals in a potential or actualized dyad. To paraphrase Festinger et al, interaction in dyads may depend more on the frequency or magnitude of the intersections of common paths than on the physical distance between primary spaces. Therefore,

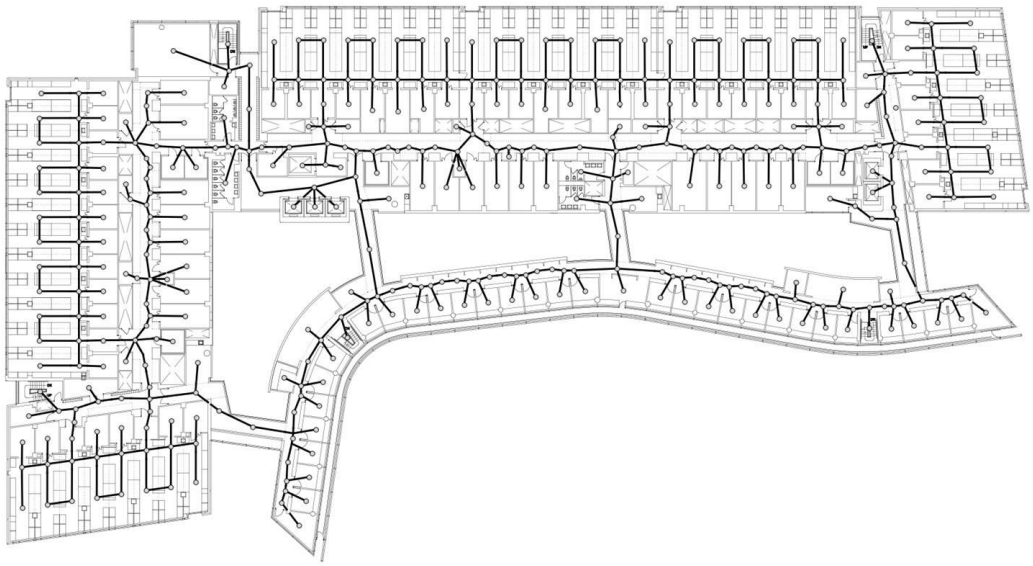
**H2**: the greater the zone overlap between two individuals the higher the potential for knowledge transfer between them and the higher their dyadic research collaboration index; this effect is robust to building layouts.

# 3. Methods

### *Participants and Research Sites*

We test our hypotheses using data from a sample of researchers working at a large public university medical school in the US in the period 2006-2010. We analyze data for researchers resident as of the end of 2006 in BLD1 (n = 166) and BLD2 (n = 94), both are biomedical research buildings which were opened or initially occupied in 2006 and 1997 respectively.

BLD1



*[Non-zero values only]*

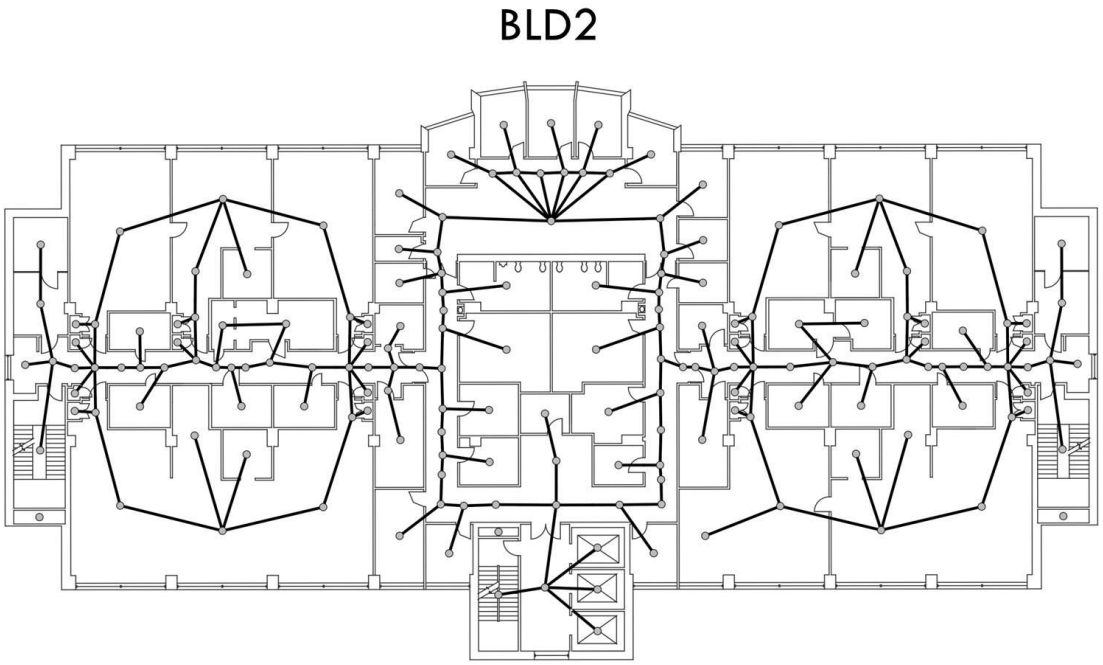
| Variable | N | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| Path overlap | 6220 | 326.97 | 219.33 | 3.8 | 1044.5 |
| Walking distance between labs | 44572 | 325.40 | 112.63 | 10.3 | 646.0 |

Path overlap & walking distance correlation = -0.563 ($p$ = 0.000)

**Figure 3.6.** Path overlap and physical walking distance at BLD1 are computed from one space to another. The image above shows the spatial network graph of one of the BLD1 floors and identifies the connections between the spaces (black lines) where spaces are connected if there is a way to physically get from one to the next.

Figures 3.6 and 3.7 show that the two buildings have different layouts; one is more linear while the other is more compact. BLD1 building has an internal atrium that separates labs and offices. The single largest contiguous part of BLD1 (the northern wing) is 428’ long by 86’ wide, giving a length-to-width ratio close to 5. In contrast, BLD2 has a compact central service core instead of an internal atrium and is 223’ long by 117’ wide giving a ratio roughly equal to 2. BLD1 is therefore more linear in its topology and longer in terms of actual physical dimensions. There was less information available on interior arrangements in BLD2, and this likely affected the granularity of the resulting spatial

network relative to the one in BLD1. The two populations were similar in terms of status and other demographic factors. Their main difference was that one of the populations moved at the beginning of the study period while the other did not. Prior to its opening in 2006, BLD1's researchers were spread out over several buildings at the medical campus.




*[Non-zero values only]*

| Variable | N | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| Path overlap | 2833 | 294.72 | 162.90 | 11.2 | 793.3 |
| Walking distance between labs | 11435 | 224.50 | 68.66 | 9.7 | 354.0 |

Path overlap & walking distance correlation= -0.548 ($p$ = 0.000)

**Figure 3.7.** Path overlap and physical walking distance at BLD2 as computed using the spatial network graph. The connections between adjacent and accessible spaces are shown (black lines).

### *Spatial Mapping*

The first step in the mapping of individuals in space is to ascertain office and lab assignments. We do this by appeal to university administrative data for regular – that is,

non-temporary – faculty for the time period 2006-2010. This data set includes human resource (HR) information on job code, department, gender, education; applications to institutional review boards (IRBs); submitted animal research protocols; successful and unsuccessful grant applications to external sponsors, and space utilization and location information (including offices and labs). To create spatial networks, we use ArcGIS and AutoCAD files for the Medical School campus in addition to finer-grained layouts for the BLD1 and the BLD2 buildings. We link the space location data to work addresses from the HR dataset to build a comprehensive picture of researchers' spatial location.

We convert electronic BLD1 and BLD2 layouts into spatial networks, decomposing the floor plan into smaller spaces as follows. First, primary assigned spaces such as offices and labs, and public and circulation spaces such as break rooms, restrooms, elevators, and stairways are treated as discrete elements. In some cases, large primary spaces are broken up into two or more subspaces so that distances between centroids accurately reflect actual walking distances. Second, connector spaces such as hallways are decomposed to identify paths between scientists' primary spaces. To achieve this goal, the connector spaces immediately adjoining the doors to primary spaces (labs and offices) are demarcated as *thresholds*. Then, connector spaces between thresholds are subdivided into smaller spaces so that the distances between the centroids of the resulting spatial element reflect actual walking distances, conditional on the arcs or edges connecting these centroids not crossing walls or other physical barriers. For typical connector spaces such as hallways, the numbers of subspaces or spatial elements between thresholds has no impact on the calculation of path and areal measures of zone overlap. That is, decomposing the hallway into many smaller subspaces versus one long space does not

change the area or path overlap. The totality of the spatial elements or subspaces constitutes a spatial network where the nodes are connected on the basis of accessibility and adjacency (see Figures 3.6 and 3.7).

### *Calculating Zone Overlap*

After generating this spatial network we map the individuals in our study, define individuals' functional zones, and calculate the zone overlaps between individuals. There are three major steps in the calculation of the zone overlap of a dyad. First, the floor plan is decomposed into spatial nodes. The distance between two nodes is computed using their centroids as a reference, provided the nodes satisfy the dual requirements of adjacency or contiguity and direct physical accessibility from one node to the other.

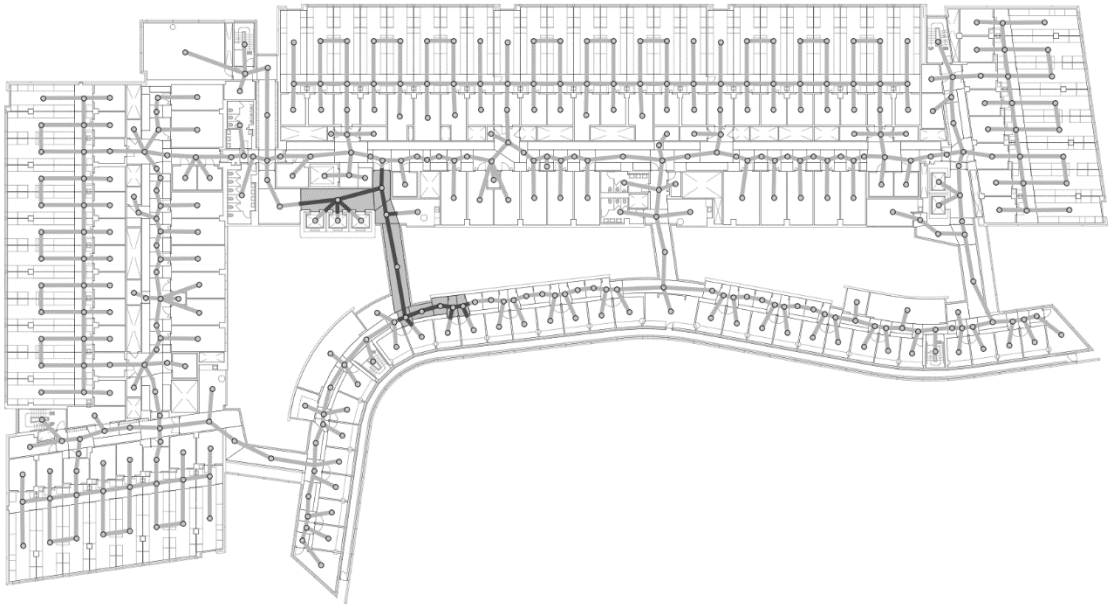


**Figure 3.8.** The zone overlap between the two individuals referenced in Figure 3.3 was computed by obtaining the intersection set of their functional zones (areal overlaps), and by summing the lengths of the paths in the intersection set (path overlap). The spaces in the intersection set are shaded in gray above while the path in shown in black.

Second, individuals' functional zones are defined. In this study zones are bounded by the following nodes: their workspaces (offices and labs), nearest public spaces (restrooms), closest circulation spaces (elevators and stairways), and all connector spaces that link them. Each person's zone is stored as a set of nodes with unique numerical identifiers. Third, the zone overlap between any pair of individuals is derived from the intersection of the sets of nodes in their respective functional zones. For example, if A's functional zone is the set of nodes [1, 3, 5, 34, 36, 45, 68, 73, 98] and B's functional zone is the set [1, 3, 5 11, 16, 25, 34, 36] then the zone overlap between them is the set [1, 3, 5, 34, 36]. We can then obtain measures of areal overlap or the sum of the area of the nodes in the intersection set, and path overlap or the sum of the total length of edges (node to node links) in the intersection (Figure 3.8).

### *Dependent Variable*

*Collaboration Index*. For any given year from 2006-2010, we create a composite index of research collaboration for each dyad in the study. This index measures the extent to which a dyad generated administrative evidence of early-stage collaboration. For each year, the index equals the sum of the following: applications to institutional review boards, animal research protocols, and grant applications to external sponsors. Because most potential dyads in the study never consummate a collaboration, the collaboration index is over-dispersed and has a left-skewed distribution.

### *Independent Variables*

*Path Overlap*. We use path overlap (measured in feet) in our regression estimates below. For these buildings, areal and path overlap are highly correlated (r = .986). Because the

interpretation of path overlap is somewhat more intuitive, we use it in our analysis. The correlations between measures of physical distance and path overlap are negative and low, suggesting that they capture complementary aspects of space (see online Appendix Tables A2 and A3). Table 1 lists the variables used in the regression models and analysis.

*Physical Distance*. We calculate three measures of physical distance: the first two are metric, walking distance ("walking") and straight line distance in feet. The third is topological, "turn". For each measure, the distance between individuals is calculated as the distance between the centroids of their primary work spaces (lab or office) using Dijkstra's algorithm (Dijkstra 1959). For individuals who had both labs and offices, we designate their primary space as the lab. Walking distance is the actual distance between these centroids, taking into account walls and other barriers as well as the presence of a physical connection between spaces. Straight line distance is the distance between centroids of spaces without consideration of barriers and physical accessibility or connections between these spaces. Turn distance is the minimum number of turns to get from one space to another. Walking and turn distances are highly correlated (r = .911).

### *Control Variables*

*Collaborativeness*. We consider a collaboration to exist whenever two people appear together on an IRB application, animal research protocol, or grant proposal to external sponsors. In order to control for personal differences in the propensity to collaborate, we create a dyad-level count variable, "collaborativeness," equal to the sum of all collaborations that each member of the dyad had with all other researchers in their building (including the other half of the dyad).

*Same Department*. Previous research has shown that affiliation, such as being in the same department, encourages and reflects homophily and, subsequently, higher levels of interaction (Agneessens and Wittek 2012; Kossinets and Watts 2006; Wineman, Kabo, and Davis 2008). We created and included a binary variable equal to one if the two people in a dyad were in the same department at any point in that particular year, and zero otherwise.

**Table 3.1.** Key Variables and Concepts

| Variable or Concept | Definition |
|---|---|
| **Collaboration index** | Yearly combination of applications to institutional review boards, animal research protocols, and grant applications to external sponsors |
| **Path overlap** | The length of the overlap in feet of the paths in the functional zones of the two people in the dyad |
| **Walking distance** | The actual distance in feet between the offices and/or labs of the two people in the dyad |
| **Turn distance** | The number of turns between the offices and/or labs of the dyad members |
| **Straight line distance** | The straight-line distance in feet between the offices and/or labs of the two people in the dyad |
| **Collaborativeness** | The sum of the number of collaborations both people in the dyad have with all other people in their respective building samples including the dyad itself |
| **Same department** | Coded as 1 if the two people in a dyad were in the same department that year |
| **Jobcode** | The variable captures whether both people in the dyad had academic or tenured/tenure-track positions (coded 0), whether one person only or half of the dyad had an academic position (coded 1), or whether both people in the dyad did not have academic positions (coded 2) |
| **Year** | The variable has a value for each of the five years in the period 2006-2010 |
| **Functional zone** | An individual's sphere of operation that is an aggregate function of the spaces that are the sites of task performance or personal movement in the workplace. In this study, there are four main types of spaces: individuals' workspaces (offices, labs), public or shared spaces (restrooms), circulation spaces (elevators, stairways), and connectors (hallways). |
| **Path overlap** | The length of the overlap – e.g. in feet – in the critical paths of two individuals' functional zones. |
| **Areal overlap** | The total area or size of the overlapping spaces – e.g. in square feet – of two individuals' functional zones. |

*Jobcode*. For academic settings the primary distinction in faculty or research appointments is between those who are tenure-track or tenured and those who are in other types of

positions. In order to account for the impact of differences in job types on collaborations in a potential dyad, we code each individual's position as "academic," (if tenured or tenure track) and "other" otherwise. At the dyadic level, we create a three-level categorical measure equal to zero if both dyad members had "academic" positions, one if exactly one dyad member had an academic position, and two if both dyad members had "other" positions.

*Year*. We use year dummy variables for each year in the period from 2006-2010.

# 4. Statistical Analysis and Model Specification

The class of Poisson regression models is best suited for count dependent variables such as our index of research collaboration. However, one of the assumptions of Poisson regression is that the mean and variance are equal. Because the dependent variable is overdispersed (the variance is greater than the mean, online Appendix, Tables A2 and A3), a more appropriate model is the negative binomial regression. Even so, a major reason for the overdispersion is the large number of zero counts for the dependent variable. The zero counts reflect that many potential dyads never form. A model that corrects for overdispersion and accounts for the large number of zero counts is the zero-inflated negative binomial regression (Karazsia and van Dulmen 2008; Long and Freese 2006). The zero-inflated negative binomial regression has two equations: a logit model to predict whether or not research collaboration occurs and a negative binomial model to predict value of the research collaboration index, given the existence of a collaboration (Long and Freese 2006).

The logit equation takes the form:

$$log(p/1-p) = \beta_0 + \beta_1 STR + \beta_2 DEP + \hat{e} \qquad \textbf{(Eq.1)}$$

Where:

| | |
|---|---|
| $p$ | = the probability of a research collaboration occurring |
| STR | = straight line distance between dyad members |
| DEP | = whether dyad members are in the same department |
| ê | = error term |

Straight line distance captures the costs or frictions of interaction between the members of a dyad in their most basic form while being in the same department proxies homophily effects as well as knowledge proximity.

The log of the research collaboration index is predicted with a linear combination of the predictor variables (Long & Freese, 2006). The negative binomial equation to be estimated is:

$$log(INDEX) = \beta_0 + \beta_1 PATH + \beta_2 DIST + \beta_3 COLL + \beta_4 DEP + \beta_5 JOB + B_6 YEAR + \hat{e} \qquad \textbf{(Eq.2)}$$

Where:

| | |
|---|---|
| INDEX | = the expected counts of the research collaboration index in a dyad |
| PATH | = path overlap between dyad members |
| DIST | = physical distance between dyad members, walking or turn |
| COLL | = total collaborativeness of the dyad members |
| DEP | = whether dyad members are in the same department |
| JOB | = job code of the dyad members |
| YEAR | = yearly fixed effects |
| ê | = error term |

We create two different zero-inflated negative binomial regression models corresponding to each of the buildings in order to account for differences in spatial layout. The first model estimates collaboration of researchers who had moved to BLD1 by the end of 2006. The second only examines researchers resident in BLD2 as of the end of 2006. In other words, even though we run models for the period 2006-2010, there are no new individuals in the two samples post-2006 (Table 3.2). However, there is attrition so the sample gets progressively smaller over time as people either leave the university or relocate to other buildings within the university. There were 4,371 BLD2 dyads in 2006 but only 2,485 by 2010. Similarly, there were 13,695 BLD1 dyads in 2006 and 8,128 dyads by 2010.

**Table 3.2.** Yearly Incidences of Dyads and Researchers

| | BLD1 | | BLD2 | |
|---|---|---|---|---|
| **Year** | **Dyads** | **Researchers** | **Dyads** | **Researchers** |
| 2001 | 3,916 | 89 | 2,145 | 66 |
| 2002 | 4,950 | 100 | 2,926 | 77 |
| 2003 | 6,670 | 116 | 3,741 | 87 |
| 2004 | 10,011 | 142 | 3,916 | 89 |
| 2005 | 12,090 | 156 | 4,371 | 94 |
| ***2006*** | ***13,695*** | ***166*** | ***4,371*** | ***94*** |
| *2007* | *11,026* | *149* | *3,486* | *84* |
| *2008* | *9,180* | *136* | *3,160* | *80* |
| *2009* | *8,646* | *132* | *2,701* | *74* |
| *2010* | *8,128* | *128* | *2,485* | *71* |

In summary, our regression models focus on inter-dyad variations while controlling for year to year variations (such as changes in NIH funding levels) that will affect all dyads in the sample. In these models a significant and positive effect of zone overlap, for instance, would suggest that collaborators with more shared pathways will have more

collaborations than collaborators with less functional overlap. We compare the effectiveness of our path overlap measure as a predictor of the dyadic collaboration index relative to walking and turn measures of physical distance for two samples of biomedical researchers in the period from 2006-2010. The two samples work in different buildings, and the differences in layout between these two spaces allow us to speak to the measure's robustness across building designs. This point is important because we believe that physical distance is susceptible to layout effects, and that any measure of functional proximity or dyadic spatial effects should be robust to layout effects in order to facilitate comparisons of different buildings.

# 5. Results

Results of the regression models are shown in Table 3.3; models 1-5 are for BLD1 and models 6-10 are for BLD2. The models were constructed as follows. First, a set of models was run with each of the three independent variables plus the control variables, that is, path overlap and controls (models 1 and 6), walking distance and controls (models 2 and 7), and turn distance and controls (models 3 and 8). Second, for each building, two sets of models were run for combinations of path overlap, either of the two physical distance variables, and controls. These were path overlap, walking distance, and controls (models 4 and 9), and path overlap, turn distance, and controls (models 5 and 10). Recall that walking and turn distances were nearly perfectly correlated and therefore were not included in the same models. Our analysis and interpretation of the results will concentrate on the full models (4 and 5 for BLD1, and 9 and 10 for BLD2).

**Table 3.3.** Effects of Path Overlap, and Walking and Turn Distances on the Dyadic Collaboration Index at BLD1 and BLD2

| *DV = COLLABORATION INDEX* | BLD1 | | | | |
|---|---|---|---|---|---|
| VARIABLES | (1) | (2) | (3) | (4) | (5) |
| **Path overlap** | 0.0021*** | | | 0.0014*** | 0.0015*** |
| | (0.0003) | | | (0.0004) | (0.0004) |
| **Walking distance** | | -0.0065*** | | -0.0044** | |
| | | (0.0013) | | (0.0015) | |
| **Turn distance** | | | -0.0593*** | | -0.0368*** |
| | | | (0.0096) | | (0.0110) |
| **Collaborativeness** | 0.0189*** | 0.0202*** | 0.0200*** | 0.0196*** | 0.0194*** |
| | (0.00125) | (0.0012) | (0.0013) | (0.0012) | (0.0012) |
| **Same department** | 1.0630*** | 0.5663+ | 0.6671* | 0.6125+ | 0.7052* |
| | (0.3095) | (0.3137) | (0.3102) | (0.3358) | (0.3301) |
| **Jobcode_Academic-Academic (reference)** | | | | | |
| jobcode_Academic-Other | -0.0904 | -0.1738 | -0.1426 | -0.1328 | -0.1088 |
| | (0.1470) | (0.1466) | (0.1483) | (0.1424) | (0.1437) |
| jobcode_Other-Other | -0.5242* | -0.6032** | -0.5510** | -0.5712** | -0.5339** |
| | (0.2048) | (0.1977) | (0.1999) | (0.1988) | (0.2004) |
| **Year_2006 (reference)** | | | | | |
| Year_2007 | 0.1917+ | 0.2104* | 0.1892+ | 0.2088* | 0.1934+ |
| | (0.1012) | (0.0981) | (0.0978) | (0.1013) | (0.1011) |
| Year_2008 | -0.1819 | -0.1887 | -0.1934 | -0.1893 | -0.1933 |
| | (0.1388) | (0.1267) | (0.1274) | (0.1315) | (0.1324) |
| Year_2009 | 0.0066 | -0.0109 | -0.0155 | -0.0015 | -0.0046 |
| | (0.1285) | (0.1205) | (0.1200) | (0.1251) | (0.1253) |
| Year_2010 | -0.4865** | -0.4676** | -0.4693** | -0.4767** | -0.4812** |
| | (0.1615) | (0.1491) | (0.1494) | (0.1571) | (0.1582) |
| Constant | -3.8368*** | -1.9821*** | -2.0943*** | -2.8373*** | -3.0075*** |
| | (0.2873) | (0.2999) | (0.3103) | (0.4043) | (0.4069) |
| *DV = COLLABORATION INDEX* | BLD2 | | | | |
| VARIABLES | (6) | (7) | (8) | (9) | (10) |
| **Path overlap** | 0.0025*** | | | 0.0018*** | 0.0026*** |
| | (0.0004) | | | (0.0005) | (0.0004) |
| **Walking distance** | | -0.0062*** | | -0.0026 | |
| | | (0.0018) | | (0.0018) | |
| **Turn distance** | | | -0.0402*** | | 0.0013 |
| | | | (0.0095) | | (0.0128) |
| **Collaborativeness** | 0.0053*** | 0.0054*** | 0.0054*** | 0.0054*** | 0.0053*** |
| | (0.0008) | (0.0008) | (0.0007) | (0.0008) | (0.0008) |
| **Same department** | 0.4959 | 0.4040 | 0.6372+ | 0.3920 | 0.5013 |
| | (0.3826) | (0.4266) | (0.3664) | (0.4199) | (0.3970) |
| **Jobcode_Academic-Academic (reference)** | | | | | |
| jobcode_Academic-Other | -0.2745 | -0.3765+ | -0.3039 | -0.3193 | -0.2730 |
| | (0.2046) | (0.2174) | (0.2048) | (0.2123) | (0.2064) |
| jobcode_Other-Other | -0.6399* | -0.8194** | -0.7278* | -0.7131* | -0.6373* |
| | (0.2834) | (0.2912) | (0.2842) | (0.2910) | (0.2890) |
| **Year_2006 (reference)** | | | | | |
| Year_2007 | 0.0775 | 0.0425 | -0.0139 | 0.0791 | 0.0778 |
| | (0.1231) | (0.1304) | (0.1216) | (0.1269) | (0.1224) |
| Year_2008 | 0.1386 | 0.1113 | 0.0896 | 0.1253 | 0.1404 |
| | (0.1462) | (0.1517) | (0.1476) | (0.1486) | (0.1449) |
| Year_2009 | -0.1273 | -0.1963 | -0.2193 | -0.1524 | -0.1250 |
| | (0.1682) | (0.1681) | (0.1661) | (0.1694) | (0.1658) |
| Year_2010 | 0.0666 | -0.0080 | 0.0122 | 0.0252 | 0.0698 |
| | (0.1561) | (0.1662) | (0.1592) | (0.1648) | (0.1576) |
| Constant | -2.4298*** | -0.9378+ | -1.3132** | -1.8562** | -2.4579*** |
| | (0.4024) | (0.4942) | (0.4818) | (0.6056) | (0.5442) |
| Observations | 44,962 | 44,962 | 44,962 | 44,962 | 44,962 |

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, +p<0.1

Across the two buildings, path overlap is significantly and positively correlated with the collaboration index even controlling for the physical distance between dyad members, thus confirming **H2**. In BLD1, a 100-foot increase in path overlap in a dyad is associated with a 14.6% increase in the expected counts of the research collaboration index when controlling for walking distance, and a 15.9% increase if controlling for turn distance (models 4 and 5). Path overlap has an even larger effect in BLD2: a 100-foot increase in path overlap correlates with 19.4% and 29.2% increases in expected counts of the research collaboration index when controlling for walking and turn distances respectively (models 9 and 10). In other words, across the two buildings, a 100-foot increase in path overlap relates to significantly higher outputs of IRB applications, animal research protocols, and grant applications to external sponsors.[7]

Physical distance is negatively and significantly related to the collaboration index in BLD1 but not in BLD2, confirming **H1a** and **H1b**. In the more linear BLD1, controlling for path overlap, increasing the walking distance by 100 feet or the turn distance by 10 turns relates to 35.4% and 30.8% decreases respectively in expected counts of the dyadic research collaboration index. In contrast, in the more compact BLD2, the correlation with physical distance is not significant when controlling for path overlap. These findings highlight the limited utility of physical distance as a proxy for the effects of spatial proximity, especially when the focal building has a more compact footprint.

The control variables performed as expected, but there are some differences between the two buildings in the “year” and “same department” variables. In both BLD1 and BLD2,

[7] In results not reported here, the logit or zero-inflation models confirm that affiliation (being in the same department) and interactions costs or frictions (straight line distance) had significant effects on the likelihood of the existence of a dyadic collaboration.

the overall “collaborativeness” of the members of a potential dyad is significantly and positively related to their dyadic research collaboration index. An increase of 10 units in dyadic “collaborativeness” relates to 21-22% and 5-6% increases in counts of the collaboration index at BLD1 and BLD2 respectively.

Being in the same department had a positive and significant correlation in BLD1 but not in BLD2. At BLD1, departmental affiliation is associated with a 84-102% increase in collaboration index counts. The “jobcode” variable shows that, despite the general trend of a steady decline in tenured and tenure-track faculty in academic institutions (Ehrenberg and Zhang 2005; Snyder and Dillow 2012), researchers fitting this description are still more likely to form or initiate new collaboration dyads relative to those in non-tenure-track positions. In BLD1 membership in all “other” dyads is correlated with 41-44% decreases in expected counts of the collaboration index. The corresponding numbers for BLD2, 47-51% decreases in expected counts, are even larger. Lastly, the year dummies are not significantly related to the collaboration index, with the exception of the year 2010 in BLD1 where there was a roughly 38% decrease in the expected collaboration index counts relative to the reference year 2006.

# 6. Discussion and Future Directions

Our analyses offer strong support for the hypotheses **H1a**, **H1b** and **H2**. Regarding **H2**, path overlap is significantly related to the research collaboration index and this correlation is similar across buildings. These effects are both substantively and statistically significant, lending credence to the utility of this dyadic spatial measure of

functional proximity. Within building micro-level differences in proximity are clearly correlated with the extent to which pairs of scientists collaborate. While more research is needed to test the zone overlap concept in other settings, our analysis suggests that a dyadic spatial measure such as ours would contribute significantly to research on relational organizational processes.

In contrast, walking and turn distances were significantly correlated with the dyadic collaboration index only in the more linear BLD1. We suspect that these differences result from the characteristics of the building layouts. We conjecture that distances matter more in BLD1 both because the occupants were relatively new to the space and it is more linear than BLD2. Despite the ease of interpretation and calculation of these measures of physical distance, their ability to capture the impact of spatial relations on social relations may be limited. One implication for future research is that there may be gains to using a tile-based computational approach to analyze more detailed layouts.[8] The larger point, however, is that more research in various types of spaces is needed to test our finding that path overlap is more strongly associated with collaboration among scientists than are walking and turn distances.

Our analysis is restricted to two buildings, making broad generalizations problematic. It is possible that the observed effects of path overlap, and walking and turn distances are due to unobserved differences in the two buildings. For example, while we attribute the differential impacts of walking and turn distances in BLD1 and BLD2 to the divergent effects of linear versus compact buildings, this could be the result of BLD1's more fine-

[8] This in turn would require efficient logics for decomposing large spaces into smaller tiles at reasonable computing costs.

grained interior detail than BDL2. While we do not think that this would demonstrably reduce the salience of zone overlap, we recognize that there are discernible differences when spatial networks are constructed with higher versus lower levels of interior detail.

We make the simplifying assumption that individuals take the shortest path available within their functional zones. While this facilitates the computation of zone overlap, we are cognizant of the limitations of this assumption. For example, the better quality coffee available in a break room farther away might make an individual take a longer path in lieu of the shorter path to the nearby break room. More importantly, an individual might forego a shorter path to avoid or see a particular researcher. Identification of the actual paths individuals take would also enable more advanced conceptualizations of functional zones. We employ a fairly simple conceptualization of functional zone; more work is needed to operationalize the different types of zones that are salient for specific workplaces.

Lastly, we acknowledge that there is endogeneity in the assignment of primary spaces. Prior relationships can influence office location and the likelihood that current encounters are related to future collaborations. The retrospective nature of our study precluded random assignment to labs and offices, raising the real possibility that there are unobserved variables that influence whether individuals collaborate and the subsequent success of those collaborations.

Future analysis of research collaboration in academic settings could address the role of departmental affiliation in fostering and maintaining collaborative efforts. Investigators are more likely to collaborate with those in their department. Whether this simply reflects common research interests or constraints on the cross-departmental collaboration

is impossible to disentangle with the data analyzed here. Moreover, our results suggest skewed relations between tenured or tenure-track, and non-tenure-track researchers. Dyads composed of tenured and tenure-track researchers are more likely to have non-zero research collaboration counts than are dyads comprised of non-tenure-track peers, who would include those with research track, clinical, adjunct, and visiting positions.

Shifting the focus of spatial analysis away from measures of distance and toward conceptions of functional zones and their overlap also suggests interesting future directions. Similarly, attention to zones and overlaps at multiple levels of analysis could shift the emphasis of both design and space allocation processes in support of research or other organizational outcomes in subtle but important ways. The most important area for future research is the examination of how people define, occupy, and traverse functional zones. We conceptualize such zones in fairly simple terms; it may be the case that other public spaces should be included in the definition. By the same token, all types of zone overlaps may not be created equal. For instance, paths that overlap as people move to and from tasks (for example, between labs and offices) may have different effects than overlaps that happen on the way to and from the restrooms or as investigators enter and leave the building. Future research could shed more light on the actual paths taken by individuals in the workplace and focus on factors that most affect path choice. A promising line of inquiry is the analysis of rich location data from wireless tracking technologies which capture the paths taken by individuals, permitting analysis of temporal and other factors in the determination of path choice.

Potentially, future research could build on earlier work that showed associations between rates of interactions and types of rooms or spaces that people occupy or move through

(Peponis et al. 2007). It would be useful for future work to elucidate how occupation of or movement through different types of spaces in the workplace engenders differential levels of awareness of others and of work activities and in turn affects the likelihood of dyadic collaborations and the probability of success for these collaborations.

It is also imperative that researchers think creatively about how to further unpack the nested effects of spatial layouts and organizational processes and structures. For example, future research would benefit from a quasi-experimental approach randomly assigning individuals to primary spaces, assuming that both were feasible and in line with broader organizational goals. Parsing the inherent endogeneity between individuals' spatial locations and organizational goals and outcomes requires more research on the link between physical space and phenomena such as collaboration.

Finally, this approach suggests new ways to consider the global impacts of small design changes. If bathrooms, for instance, are important markers of functional zones, then buildings that separate male and female facilities at opposite ends of long hallways will systematically increase zone overlaps between same sex pairs while diminishing them for mixed sex pairs. In that case, our findings strongly suggest that such a design will increase rates of same sex collaborations while decreasing the incidence of mixed sex collaborations. This possibility hints at some of the subtle mechanisms by which decisions about the design and allocation of space serve to create, sustain, or ameliorate significant workplace differentials. Conceptualizing and measuring proximity effects in terms of flexible, overlapping zones of activity that take into account the contingent ways individuals occupy and make their way through buildings offers new possibilities for research that advances theory while having immediate relevance for policy and design.

CHAPTER 4

# Network Communities in the Visibility Graph:

## A new method for the discretization of space†

## 1. Introduction

Translating space into a discrete system by decomposing the space into subspaces benefits to researchers. Firstly, representing space as a discrete system helps researchers to easily understand a spatial structure by decreasing its complexity. Secondly, once a system becomes discrete, analytic methods that are only applicable to discrete systems–for example, network analysis–become available to apply. The benefit of describing space as a discrete system in terms of analytic power has been well established by Space Syntax over a few decades. However, to translate a spatial layout into a discrete system is not an easy task (Peponis and Wineman 2002). Only a few rigorous methods for the discretization of space have been developed so far such as convex partitioning (Hillier

† This paper was originally presented at *the 9th International Space Syntax Symposium 2013*. Hwang, Y. 2013. "Network Communities in the Visibility Graph: a New Method for the Discretization of Space." In *Proceedings of the 7th International Space Syntax Symposium*. Seoul.

and Hanson 1984) and e-partitioning (Peponis et al. 1997), both of which have limitations.

Firstly, current methods require a certain degree of approximation in spatial structure or minor violations of protocols when the methods are applied to spaces with, for example, a curved wall, a concave wall, a free-standing column, and a small indent. This looks trivial when such elements are negligible. However, this becomes a source of arbitrariness in decomposition when those elements are no longer negligible.

Secondly, current methods consider only how well a 'cut' of space achieves the internal completeness of the resultant subspace such as convexity. The methods do not consider how well the cut *separates* adjacent subspaces. For example, what convex partitioning cares about is whether subspaces resulting from the cut are convex, not how effectively the cut separates the two subspaces. In addition, current methods do not utilize the global property of spatial structure during the partitioning process; instead, only local spatial structure is used in determining the location of a cut.

Lastly, current methods do not have a well-defined way to adjust the 'resolution' of decomposition. A researcher working on an analysis of a huge building may want to ignore trivial violations of the convexity rule in convex partitioning to reduce unnecessary bias in the building's spatial configuration. Or the researcher may want to merge small e-spaces with tiny differences in visual information to reduce unnecessary complexity beyond one's purpose of analysis. However, there is no 'native' method or well-established process for current methods to adjust the coarseness or fineness of decomposition. What a researcher can usually do is ignore trivial violations or merge small subspaces on the fly at the risk of arbitrariness.

The aim of this study is to propose a new method for space discretization that is readily applicable to the real world's space with curved walls, columns, and indents; that provides a partitioning process considering the quality of spatial separation with a view to global spatial structure; and that enables researchers to adjust the level of analytic resolution within its process. The core idea of this method is to apply a community detection algorithm to the visibility graph of a floorplan so that a 'community' of grid points in the visibility graph becomes each partitioned space.

# 2. Background

## 2.1. Need of area-based decomposition

There have been three ways to translate a floorplan into a discrete system: line-based decomposition, area-based decomposition, and point-based decomposition. Each approach has its own strengths and weaknesses.

Line-based decomposition usually means axial line decomposition. Axial line representation of space is suitable for linear spaces such as streets or roads when a researcher wants to define distance as the number of turns or the angle of turns instead of metric distance. For indoor spaces, it is fairly useful for closed plans with well-defined corridor spaces and cell-type offices. However, axial line decomposition is often too coarse for indoor spaces, and it does not appropriately represent the spatial structure of open plans.

Since the Visibility Graph Analysis (VGA) technique was introduced (Turner and Penn 1999), area-based analyses such as convex map analysis are sometimes considered as an inferior alternative to point-based analyses such as VGA. VGA has two strengths over area-based analysis. Firstly, it has much higher 'resolution' of analysis and hence it is suitable for very detailed analyses. Secondly, it requires much less effort to build a spatial network than convex partitioning. Unlike convex partitioning that requires researchers to manually draw convex polygons and link them in order to build a spatial network, VGA has an automated process of building a spatial network.

The visibility graph analysis, however, tends to over-estimate a large room's spatial properties under some circumstances. This mostly happens when a researcher takes the average property of points in the space as the property of the space, which is a very common practice. Let's think about an imaginary floor with two rooms connected by a corridor space (see Figure 4.1). The floor has a symmetrical layout except for the size of the two rooms. We are interested in the mean depth of each room. Let's build a visibility graph on the grid points in the floor. From a grid point in a room, all of the grid points in the same room are one step away in terms of visibility; the grid points in the corridor are two steps away; finally, the grid points in the other room are three steps away, which is the maximum distance. Then, the average of the mean depth of the grid points in the large room would be significantly smaller than that of the grid points in the small room, because a grid point in the large room has more grid points at one step away and less grid points at two step away than a grid point in the small room. In Figure 4.1, for example, from point A in the large room, 29 points are at one step away, 10 points at two steps away, and 10 points at three steps away. So the mean depth of point A is 1.61. All points

in the large room will have this value if we ignore trivial exceptions on the border. From Point B in the small room, 10 points are at one step away, 10 points at two steps away, and 29 points at three steps away. So the mean depth of point B is 2.39 and all points in the small room will have this value.

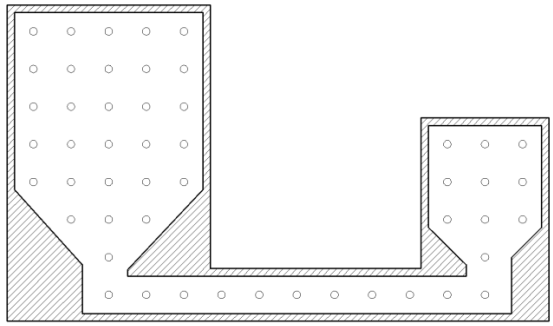


**Figure 4.1.** Imaginary floor with two rooms. The floor has a symmetrical layout except for the size of the two rooms.

Thus, if we take the average of the mean depth of the grid points in each room as the representative value of mean depth, or accessibility, of each room, then the large room becomes much more accessible than the small room even though the two rooms occupy symmetrical locations. Such bias happens solely because the large room is large and small room is small, not because the two rooms occupy such spatial locations.[9] And the bias would not happen if we decomposed the floor into three 'areas', instead of 49 'points'.

[9] The bias would not go away when we calculate mean depth without the points in the same room. The bias persists because the size of the other room matters. For example, if we calculate the mean depth of a point in the large room excluding the points in the same room, the mean depth is (10*1 + 10*2)/20 = 1.5. And the mean depth of a point in the small room is (10*1 + 29*2)/39 = 1.74. Still, the large room has smaller mean depth.

Area-based decomposition is also useful when we need to count spatial events. For example, we often need to answer a question like "which space is the most frequently used for X?" or "where did X happen the most frequently?" because of very practical reasons. Also, by describing spatial events as frequency, a researcher can use statistical techniques for count variables. In order to count such frequencies, we need a 'bin' to put events in. Area-based decomposition can provide appropriate bins for such purpose.

## 2.2. Traditional ways of area-based decomposition

### 2.2.1. Convex partitioning

Currently, the most widely used rigorous methods for decomposition of indoor space is convex partitioning. This method aims to decompose a floor into the fewest number of convex polygons required to cover the floor (Hillier and Hanson 1984). Partitioning space into convex polygons is a very powerful concept because the convexity of space guarantees that any two people in the space can see each other. However, the process of the convex decomposition is difficult to automate because it is known as an NP-complete problem[10] (Karp 1972). Thus convex partitioning has been mostly conducted manually.

Manual convex decomposition is vulnerable to the arbitrariness of an operator who decomposes a floor because it is almost impossible to strictly apply the convexity rule under some circumstances.

[10] Convex partitioning is equivalent to finding the minimum clique cover of the visibility graph of the floor, one of the well-known graph theoretical problems because a convex space is equivalent to a maximal clique in the visibility graph. The minimum clique cover problem is known to be NP-complete.

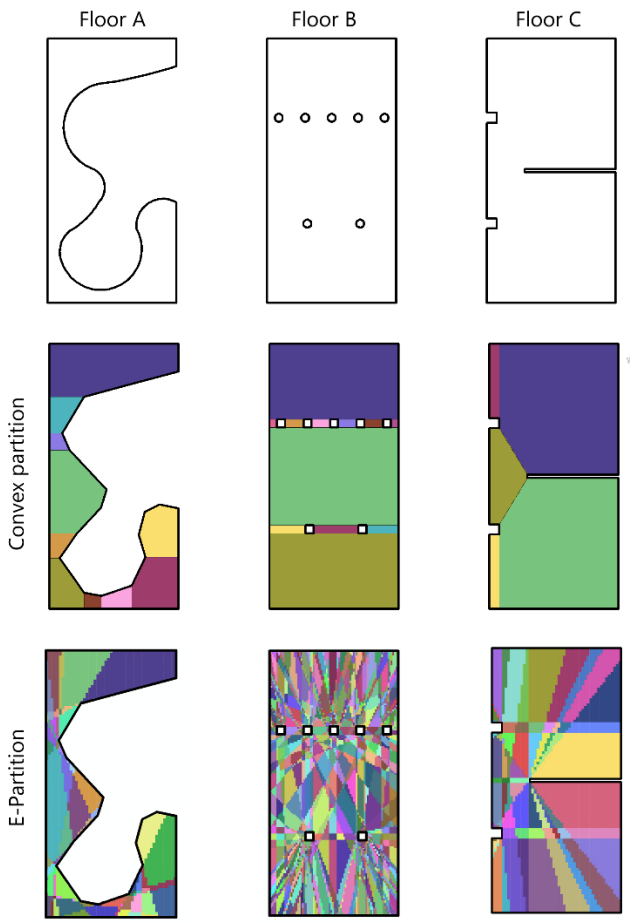



**Figure 4.2.** Exemplar floors for traditional methods. Subspaces generated from traditional methods such as convex partitioning and e-partitioning are illustrated.

Firstly, convex decomposition is not applicable at all to space with a concave curved wall like Floor A in Figure 4.2. To decompose such a floor to convex spaces, we have to approximate the concave curved wall with a series of straight walls. Thus the composition of space is mainly determined by the way the curved wall is approximated with straight walls. This means a considerable amount of arbitrariness might be involved in the decomposition process.

Secondly, convex decomposition does not handle well a space with small ‘indents’ or ‘bulges’. For example, the upper room on Floor C in Figure 4.2 would have only one convex space if there were no indents. However, because of the indents, the room should be split into three convex spaces lest the convexity rule be violated. If the indents become smaller, we would be more tempted to ignore such indents and to violate the convexity rule. This is another source of arbitrariness.

Thirdly, a more complicated case happens when we try to decompose space with a column. Let’s assume that we are going to decompose a room with free standing columns like Floor B in Figure 4.2 using the convex partitioning method. There are three options. The first option is to literally follow the convexity rule and to get 12 convex spaces[11] from the simple room. The second option is to simply ignore all columns and to get only one convex space as if the columns are nothing to do with the separation of the space. The last option is to ignore the columns’ size without forgetting the existence of the columns. This option gives us two or three convex spaces depending on the threshold for determining how many consecutive columns are regarded as separators of space. If we set the threshold at four columns, we would have two convex spaces; if we set the threshold at two columns, we would have three convex spaces. Such a threshold is, again, mostly arbitrary.

We may think we can address the problem of arbitrariness by listing ‘exception rules’ related to convexity condition such as to ignore small indents less than one meter’s offset; to straighten concave walls with one meter segments; to ignore the thickness of a wall

[11] If the columns are circular ones, the columns should be approximated to straight lines before we apply the convexity rule. We get 12 convex spaces when the columns are approximated to a rectangular shape. If the columns are approximated to an octagon, many more convex partitions would be required.

less than 0.3 meter; to approximate a circular column as a rectangular column; to ignore the spaces between columns closer than one meter and so on. However, to list a complete set of exception rules seems not an easy task, and the rules may conflict with each other. Moreover, to translate such rules into machine-understandable language looks even more difficult. Partly because of the lack of a well-defined set of machine-understandable exception rules, and partly because of the inherent computational complexity of convex partitioning, there is no widely used computer program for the automation of convex decomposition to the best of our knowledge. Thus convex partitioning is still conducted manually and remains vulnerable to operational arbitrariness.

### 2.2.2. E-Partition

E-partition is a method of dividing spaces into e-spaces that are 'informationally stable' in the sense that any location in the e-space shares the same set of visible vantage points such as corners or end points of walls (Peponis et al. 1997). This has been the most rigorous approach to understand the morphology of floor plans based on visual information. The method provides a mathematically well-defined set of subspaces, and its computational complexity is significantly lower than that of convex partitioning.

In spite of its theoretical elegance, the e-partition method has some limitations in its application. Firstly, the method is difficult to apply to a floor with curved walls or boundaries that have no salient vantage points. In such a case, we need a rule for designating a point on the curve as a vantage point or a rule for converting the curve into straight line segments as we did for Floor A in Figure 4.2.

Secondly, in many practical cases, the e-partition method generates a large number of tiny e-spaces usually unsuitable for common analytic purposes. This is because the number of e-spaces dramatically increases and hence the size of the e-space rapidly shrinks as the complexity of space increases[12]. Although having many small partitions helps a researcher to conduct a fine-grained analysis, too many partitions make the floor's spatial structure almost illegible (see e-partitions of Floor B in Figure 4.2).

# 3. Proposed Method

The core idea of this method is to decompose the visibility graph of a floorplan into closely interconnected groups of nodes. Then, the key becomes how to find such groups in the network and how appropriate the grouping would be for various applications. For this aim, this study utilizes a community detection technique developed for discovering closely related groups in a network. Later on, the new method proposed in this study will be called the NCVG (Network Communities in the Visibility Graph) method.

## 3.1. Visibility graph

Unlike convex partitioning or e-partitioning that 'cuts' a floorplate into polygons with dividing lines to make subspaces, this method defines a subspace with a set of grid points on the floorplate in the same way that an isovist from a point can be defined as a set of points that are visible from a focal point. This approach that defines an area with a point

[12] In general, the number of e-spaces is proportional to the quadruple of the number of vantage points. The number of 'cutting edges' is proportional to the square of the number of vantage points, and the number of sliced planes (e-spaces) is proportional to the square of the number of cutting edges.

set is also applicable to traditional methods like convex partitioning or e-partitioning. For example, a convex space can be defined as a set of grid points that forms a clique, or a complete graph. An e-space can be defined as a set of grid points sharing the same set of visible vantage points. As seen in the point-set based representations of a convex space and an e-space, the key process to identify a subspace is to identify a group of grid points that are closely related or that share similar attributes. This method finds subspaces by identifying closely related groups of grid points using network structure of the visibility graph of a floor.

## 3.2. Community detection

Figure 4.3 shows a network structure with three obvious groups of nodes in the network. Each group has relatively many edges inside the group and fewer edges going outside of the group. Such a group is often called a 'community' by network scientist, which is usually defined as "a cohesive group of nodes that are connected 'more densely' to each other than to the nodes in other communities"(Porter, Onnela, and Mucha 2009). Finding a community structure of a network has proved to be useful in many fields. In biology, for example, community detection algorithms have been applied to protein-protein interaction networks to identify functional modules of proteins (Chen and Yuan 2006). In communication networks, by analyzing the community structure of an email exchange network in scientific labs, researchers could identify groups of people quite closely matched to the labs' organizational structure and project assignment (Tyler, Wilkinson, and Huberman 2005). An analysis on the community structure of a citation network over

600 scientific journals presented a ‘map of science’ showing how each discipline in science is related to others (Rosvall and Bergstrom 2008).

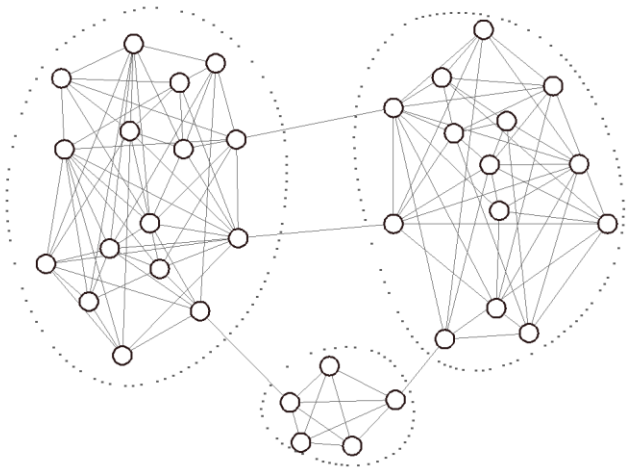


**Figure 4.3.** A network with obvious community structure. Each group has relatively many edges inside the group and fewer edges going outside of the group. Such a group is often called a ‘community’.

There is also a traditional way of decomposing a network into subcomponents, which is often called graph partitioning. Unlike the community detection methods, graph partitioning requires to fix the number of clusters to be separated in advance (Newman 2006). The main goal of graph partitioning is to find the best division of a network with the given number of divisions; hence, this approach is useful when we have a strong reason for the pre-fixed number of divisions. In this paper, we will not follow this line of approach as we do not have reason to specify the number of clusters in general spatial decomposition. So we will place more focus on community detection techniques, which does not require identifying the number of clusters in advance.

### 3.2.1. Modularity

The crux of community detection is how to find 'good' partitions of a network. One approach is to set up a quality function measuring how good current partitioning is and then to optimize the quality function. Currently, one of the most widely used quality functions is Newman's *modularity* Q (Newman and Girvan 2004):

$$\mathrm{Q} = \frac{1}{2m}\sum_{ij}\left[A_{ij} - \frac{k_i k_j}{2m}\right]\delta(c_i, c_j) \qquad \text{(Eq 4.1)}$$

where $m$ is the total number of edges in the network, $A$ is the adjacency matrix of the network, $k_i$ is degree of vertex $i$, $\delta(c_i, c_j)$ is the delta function whose value is 1 when vertex $i$ and $j$ belong to the same cluster and 0 otherwise. In plain words, the modularity function Q measures the difference between the existing number of edges in the cluster and the expected number of edges in the cluster when the network is randomly wired ignoring community structure (Newman 2006). Thus, maximizing modularity Q means minimizing the difference between the number of intra-group edges and the expected number of inter-group edges (Fortunato 2010).

A critical difference between modularity based decomposition and convex partitioning is that convex partitioning does not consider how inter-group edges are distributed across groups. Look at Figure 4.4 showing a floor with three rooms connected with two openings; one is wide and another is narrow. Without any complex analysis, we can say that room A and room B form almost one space, whereas room B and room C are quite separated. Convex partitioning ignores such an obvious difference between the two openings and yields three convex spaces (if wall thickness is ignored).

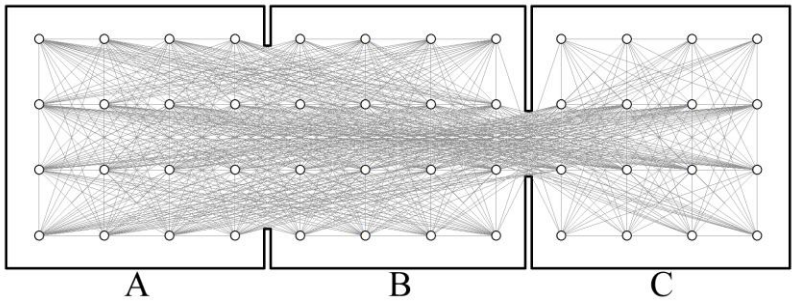



**Figure 4.4.** Floor with two thresholds: one is wide and another is narrow. Room A and room B form almost one space, whereas room B and room C are quite separated. The visibility graph of the floor illustrates the difference between the two openings.

The difference between the two openings becomes clearer when we construct a visibility graph on the floorplan in Figure 4.4. There are much more inter-edges passing the opening between A and B (total 311 edges) than inter-edges passing the opening between B and C (total 181 edges). Thus, modularity based partitioning prioritizes B-C cut to A-B cut, whereas no priority is possible with convex partitioning. In other words, the modularity based partitioning does not guarantee the 'completeness' of connections as convex portioning does; however, it helps to find the most 'effective' cut maximizing internal connections while minimizing external connections.

Another noteworthy difference in modularity based decomposition is that it gives a decomposition based on the global property of a spatial network, while convex partition or e-partition uses only the local property of the space. For example, whether a space is convex or not has nothing to do with other parts of a floor. It can be determined only with local information. In contrast, modularity-based decomposition cannot be done without the knowledge of the whole network structure. For example, let's assume that the rooms in Figure 4.4 are only a part of a large floor.  Then, whether there would be any partition

inside A-B-C cannot be determined without information on other spaces connected to A-B-C. If there are plenty of good locations for 'cuts', then we would be less likely to have a partition within A-B-C. On the contrary, if outside spaces are all very well-connected, then we would be more likely to have a partition within A-B-C.

### 3.2.2. VOS clustering technique

Among many community detection algorithms proposed so far by network scientists, this paper uses VOS[13] clustering technique(Van Eck and Waltman 2007) for finding communities in space. This is because the method can adjust 'resolution' of analysis, respects edge weights, and has reasonable computation cost. The VOS technique has a slightly different form of quality function from Newman's modularity Q in order to introduce a resolution parameter γ and the degree of association strength between nodes (edge weight). The quality function V of the VOS technique is:

$$\mathrm{V} = \frac{1}{2m}\sum_{ij}\left[s_{ij} - \gamma\right]\delta(c_i, c_j) \qquad \text{(Eq. 2)}$$

where $s_{ij}$ denotes association strength between vertex $i$ and $j$, which is given by

$$s_{ij} = \frac{2mk_{ij}}{k_i k_j} \qquad \text{(Eq. 3)}$$

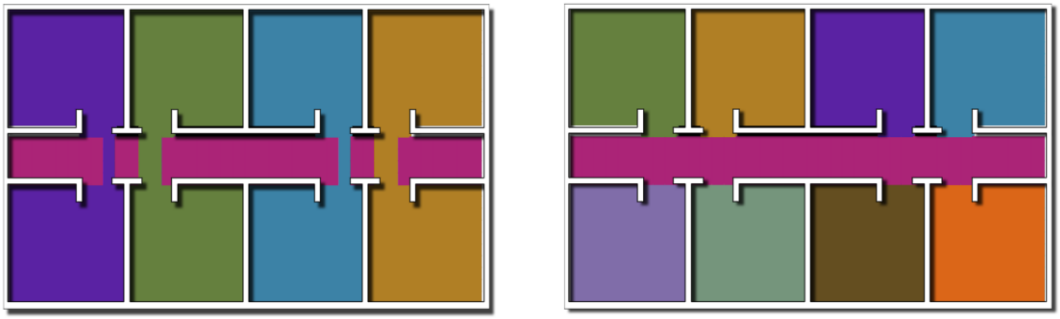
where $k_{ij}$ denotes the number of edges between vertex $i$ and $j$.

[13] VOS stands for visualization of similarities

It has been shown that the quality function V of the VOS clustering technique is equivalent to Newman's modularity Q when resolution parameter γ and edge weights are set to one (Waltman, van Eck, and Noyons 2010).

## 3.3. Distance weight

Figure 4.5 shows a sample result of VOS partitioning (γ=0.4) applied to a typical floor with a corridor. We can see the upper rooms and the lower rooms are pairwisely interconnected. This is an understandable result, although not desirable, with the given visibility graph structure that connects any visible grid points *equally* no matter how near or far they are. In other words, the visibility graph does not contain information about the 'locality'. Hence, a distant pair of points across the corridor can be equally grouped together because they are visually 'neighbors' as a closely located pair of points within the same room.  This is not what we expected to obtain from floorplan decomposition.



(a) No distance weight (b) Inversely distance weighted

**Figure 4.5.** Comparison between no weight visibility graph (Panel A) and distance weighted visibility graph (Panel B)

The problem can be solved by emphasizing the locality of a visibility graph. In other words, one way of to strengthen locality is to make ties between nearer points stronger

and to make the ties between farther points weaker. Thus, we give edge weights to the visibility graph according to the inverse of the distance between the two points. Panel (b) in Figure 4.6 is the result of the same algorithm as panel (a) except for the edge weights. It results in a set of partitions that fit with our intuitive partitioning.

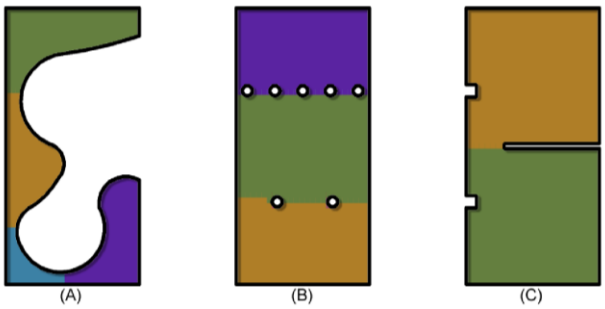



**Figure 4.6.** Application of the NCVG to the three previous floors. The results coincide with intuitive decomposition.

## 3.4. Implementation

We developed software for calculating and visualizing the NCVG method. ArcPy, the official Python package for ArcGIS by ESRI, was used for generating grid points, checking visibility, and visualizing the decomposition result. NetworkX, an open-source Python package for network analysis was used for general network operation such as constructing a visibility graph, giving edge weights, and exporting a graph. VOS clustering software by van Eck and Waltman was used for the calculation of the VOS algorithm. Python was used for a 'glue' language combining these components.

The software gets shapefiles of a floor's boundary and internal walls as inputs. Then it builds a weighted visibility graph and determines groups of nodes in the visibility graph

according to the given resolution γ. The outputs are a shapefile containing grid points with partition information and an image file visualizing the partitions.

# 4. Discussion

## 4.1. Revisiting the three floors

Figure 4.7 shows a result obtained by applying the NCVG method (γ =0.4) to the three floorplans previously shown in Figure 4.2. Floor A with a curved wall is decomposed into four subspaces being separated at "bottleneck" locations. Floor B with seven free standing columns is successfully separated into three subspaces. In floor C, the algorithm ignores small indents in each room and identifies two rooms following the designer's intent.

One might ask, "Can we have more (or less) divisions for Floor A?" Or, "The upper part of Floor B is separated by five columns, while the lower is separated only by two columns. Should we have to treat them equally?" Or, "What would happen if the indents in Floor C become larger? Do we still have two subspaces?" These questions are closely related to the effect of the resolution parameter.

## 4.2. Effect of resolution

One of the major advantages of the NCVG method is that we can control the resolution of analysis. For example, if we need to conduct a fine-grained analysis of a floor, we would use a higher resolution parameter that identifies more groups. Let's look at an exemplar floorplan. How many subspaces would you expect to see from the floor (a) in Figure 4.7?

The panels from (a) to (e) in Figure 4.7 show how subspaces are differentiated by the level of resolution. If a researcher wants to see a big picture or to identify the most critical 'cuts', he or she would set the resolution low enough and get two subspaces as in Figure 4.7-(b), saying, "What is important in this spatial structure is the long bottleneck in the left part, and we should ignore the trivial indents in the right part." Or, he or she would choose panel (e) if he or she thinks, "There are three rooms and two connectors in this space. They are all independent spaces and I'd like to see them all."

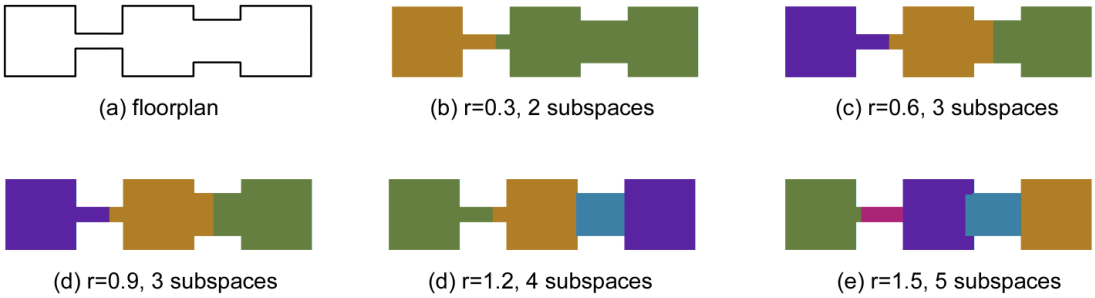



**Figure 4.7.** The effect of the resolution. The panels from (a) to (e) show how subspaces are differentiated by the level of resolution.

Going back to Figure 4.6, we can predict how the three floors will be decomposed as resolution changes. For floor A, we will have more divisions as the resolution parameter increases. Floor B will have only one subspace with very low resolution. As the resolution parameter increases, the upper part of Floor B will be identified firstly and then the lower part will be also identified. The small indents in Floor C will be ignored when the resolution parameter is small. However, when the resolution increases or the

size of indents becomes larger, the floor will be divided into more than two subspaces because the indents are not to be ignored.

## 4.3. Hierarchical spatial structure

One useful aspect of the series of decomposition with changing resolution is to reveal underlying spatial structure. Figure 4.8 shows the series of spatial differentiations from a simplified Miesian house. With a low resolution parameter ($\gamma$ =0.10), the floor is divided into the three subspaces: MNOP – QR – STU. As the resolution $\gamma$ increases to 0.30, the largest part MNOP is differentiated into three subspaces M-OP-N. At $\gamma$ = 0.40, STU cluster is divided into S and TU. At $\gamma$ = 0.80, OP is divided into two subspaces, O and P. At $\gamma$ =1.20, Q becomes independent from R. Finally, T and U are differentiated at $\gamma$ =1.40.

As more spaces become differentiated, the underlying structure of the floorplan becomes clearer. As we see in the overlaid graph representation in Figure 4.8, the floorplan is composed of the main ring structure in the left part (M, N, O, and P) and the linear tail structure in the right part (N, S, T, and U). This structure is not emergent at a low resolution such as panel (a). At $\gamma$ = 0.30, the ring is identified, and the tail becomes longer at $\gamma$ = 0.40. The ring structure is fragmented into five subspaces at $\gamma$ = 0.80, and a short branch (Q) attached to the ring is identified at $\gamma$ = 1.20. At $\gamma$ = 1.40, the tail is further differentiated.

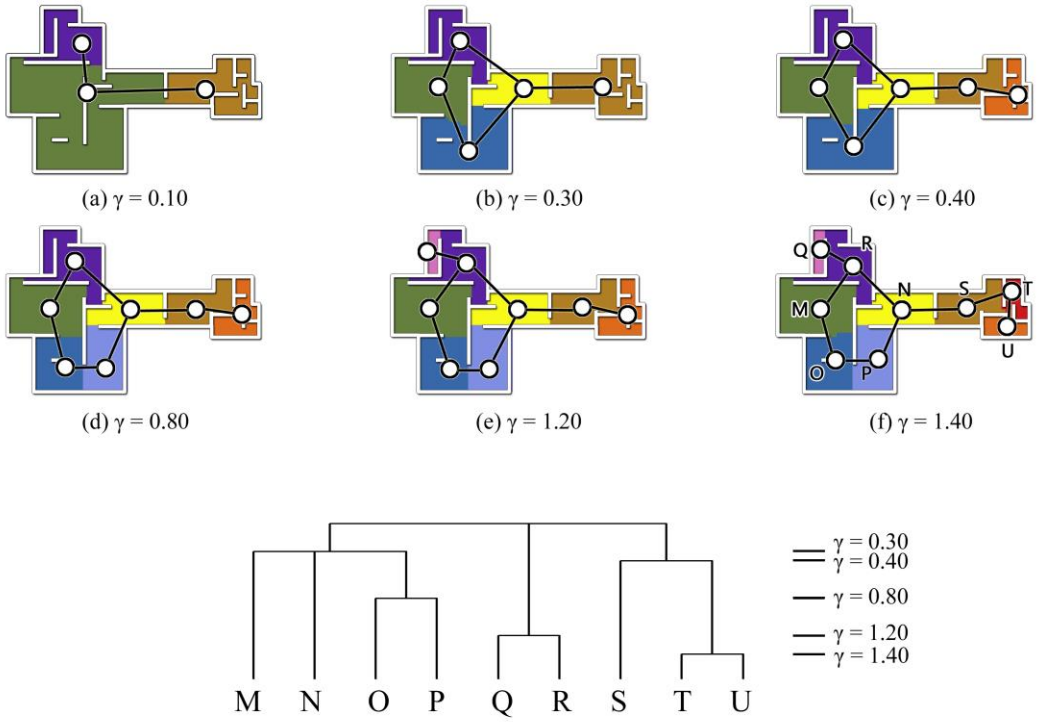



**Figure 4.8.** Differentiation of Space. Changing resolution is to reveal underlying spatial structure.

The differentiation process is summarized as a dendrogram in Figure 4.8. The dendrogram gives a picture on the hierarchical spatial structure of the floor, which is not easily captured in the traditional 'snap-shot' graph representation of a floorplan. Space S is adjacent both to N and TU; however, the S-N connection and the S-TU connection are not equal. S is more closely related to TU than N in that S-TU is merged together earlier than S-N. Likewise, space P that is adjacent both to N and O is merged first with O and then merged with N, which means the P-O connection is stronger than the P-N connection.

# 5. Conclusion: limitations and further work

In this paper, we present a new method for the decomposition of space requiring no ad-hoc rules and relatively free from an operator's arbitrariness. Unlike traditional methods, this method decomposes space based on the global property of a spatial network by adopting modularity function as a quality function of decomposition. The VOS technique used for this method enables researchers to adjust the level of analytic resolution. This gives much more flexibility in the analysis of space. Also, this method provides an analytic tool for exploring spatial hierarchy by the examination of the series of spatial differentiations.

Although this paper provides an interesting approach to spatial decomposition, much work is still required.

First of all, it will be interesting if the result of the decomposition using this method is compared to human recognition of space or actual space utilization. For example, we may compare the decomposition of an office floor to the work groups in the organization occupying the floor. If the floor is used by 5 groups in the organization, we may decompose the floor into several spaces by choosing appropriate resolution $\gamma$. Some work group's territory would be well-matched to the decomposition result while other group's territory would be poorly-matched. Then, for each group, we might compare the degree of matching to the group's interaction patterns with other groups or the group's sense of belonging.

Secondly, a dedicated algorithm for community detection in the spatial network may need to be developed. Currently we are using the VOS algorithm developed originally for

citation networks, not for spatial networks. We may expect better performance and more control on the community detection process if we have a tailored algorithm for spatial networks. Thirdly, applying this method to three dimensional space might be interesting. One of the strength of this method is that this method can be expanded to a three dimensional environment with almost no additional effort once the three dimensional visibility graph is given.

Finally, we may explore the possibility of applying community detection techniques to spatial networks other than visibility graphs. For example, we may apply community detection techniques to the road network of a city in order to identify 'natural' urban tissues of the city. Then we may explore how such identified urban tissues are well or poorly matched to actual urban behaviors or perceived urban images. For example, Figure 4.9 shows the decomposition of Ann Arbor's street network at several levels of the resolution. We may compare the decomposition result to a resident's perceived boundary of his/her neighborhood, or a resident's social network.

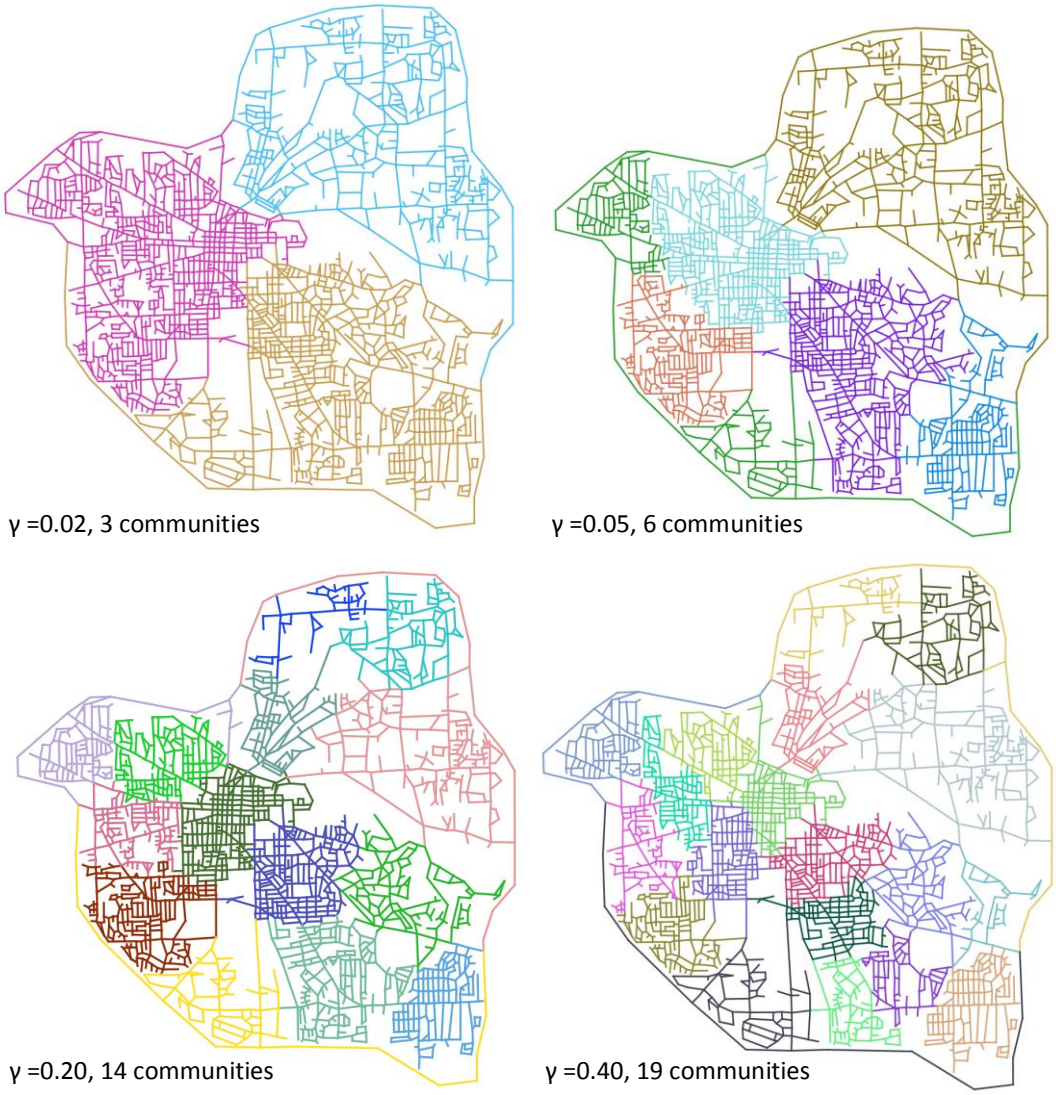



Figure 4.9. Decomposition of Ann Arbor's Street Network

# CHAPTER 5
# Conclusion

In this chapter, we summarize the major contributions of the dissertation and how it advances current space-communication studies. This chapter also suggests the implications of the result of this dissertation to design practice. Lastly, the chapter discusses several limitations of the research and the direction of further research work. This chapter concentrates on addressing common conclusions across the three essays in this dissertation. See the conclusions in each essay for the contributions, limitations, and future directions specific to each essay.

## 5.1. Dissertation Contributions

The findings from this study make several contributions to the current literature. First, the dissertation provides new metrics for two different categories of the current field of space-communication studies. As reviewed in Chapter 1, the current studies in this field can be classified into three categories: location-level studies, path-level studies, and layout-level studies. The two essays in this dissertation propose and validate two novel metrics, *sociospatial betweenness* and *zone overlap*, for the first two categories.

The two metrics proposed in this dissertation are fundamentally different from the traditional metrics used so far. Traditional metrics usually aim to measure accessibility of a space or proximity between spaces, both of which emphasize distances between spaces. In contrast, the new metrics aim to measure spatial properties that have been rarely investigated such as confluence and exposure, both of which emphasize the likelihood of encounter among people.

For the two metrics newly proposed in this study, we demonstrated their validity with empirical datasets, and compared them to their traditional counterparts, metrics for accessibility and proximity. In Chapter 2, the empirical validity of sociospatial betweenness was demonstrated through the analysis of a floor occupied by a division of a manufacturing company. Sociospatial betweenness of a space was found to be positively associated with the diversity of communication partners; in contrast, traditional spatial betweenness did not show such an association. In Chapter 3, from the analyses of two research buildings with very different layouts, we found that increasing path overlap is associated with increases in collaborations in both buildings. In contrast, traditional metrics influence outcome measures in only one of the research buildings.

One possible reason of the new measures' better associations to communication behaviors is that both of the metrics actively utilize non-spatial information, and integrate such information to the spatial structure. For example, a functional zone of a person cannot be demarcated without information on what the occupying organization's key attractors and facilities are and where they are located in the building. Also, sociospatial betweenness requires information on the locations of the building occupants and the social structure among them. The essays provide examples illustrating how spatial

metrics can be enhanced by embracing organizational and social information and how such attempts are successful in contrast to traditional metrics utilizing only spatial information.

In Chapter 4, we proposed a new method for spatial decomposition, and demonstrated that the new method successfully addresses the problems of traditional methods. Although spatial decomposition is one of the essential processes for the analysis of building layout, no new rigorous decomposition has been proposed until this study for more than a decade. The essay introduced the modularity function as a quality function to evaluate the goodness of spatial decomposition. Previous decomposition methods so far have rarely paid any attention to the evaluation of decomposition. In addition, the method saves a significant amount of manual processing time by automating the decomposition process with adjustable resolution.

## 5.2. Design Implications

The findings of this study have a number of important implications for future practice both on space management and the design of space.

First, the results of this dissertation indicate that facility managers can use common spaces as a tool for regulating the current patterns of movement and interaction in the workplace in a more rigorous manner. As discussed in Chapter 3, an employee's functional zone for daily activities changes according to the locations of key common spaces, and the patterns of overlap of such zones affect how much people encounter each other. Thus facility managers can adjust the range of movement and the degree of unplanned encounters by their choice of key facilities' locations. This is not a new

discovery as the anecdote about Steve Jobs' radical removal of restrooms shows (see Chapter 1). Advances proposed by this study are that facility managers can make such design decisions with evidence and that they can make a quantitative prediction of the effect of their design decisions. For example, when a facility manager needs to choose the location of a new coffee bar, the manager may compare how much overlap would increase in general by choosing one location among several alternatives, and which units' overlap is most increased by the manager's choice. Thus, if the organization wants to promote collaboration across a specific pair of units in the organization, the facility manager may choose the location of the new coffee bar maximizing the pair's overlap.

Second, the methods proposed in this study help a designer to predict core locations in terms of interaction activities. We developed a method predicting the degree of confluence of movement for a space. A space of high confluence would be a space used by more diverse users as we demonstrated in Chapter 2. Thus, if a designer wants a facility to be used by many different people like the lobby of Zappos discussed in Chapter 1, it should be located at a location of high confluence, which can be measured by sociospatial betweenness. In addition, a designer may overlap all employees' functional zones to identify spaces used by many different people. If the locations of activity cores or interaction cores revealed by such methods are different from the designer's intention, the designer may want to adjust the locations of key facilities or the corridor structure to match design intention with predicted use.

Finally, the findings of this dissertation urge designers and facility managers to pay more attention to non-distance-based properties. Our findings suggest that the overlap of movement paths is also a critical factor affecting communication patterns in

organizations as well as walking distance. In the future, designers might be asked to design a corporate campus with highly overlapped functional zones, as the designers of the Google campus were asked to design the campus with smaller walking distance. Also, facility managers might be asked to adjust the locations of common spaces to maximize confluences, as with the shutdown of restrooms in Pixar studio.

## 5.3. Limitations and Future Directions

***Layout level analysis***

Further research might explore the methods of the layout-level analysis. In this research, we proposed and validated new methods for the two categories – path-level analysis and location-level analysis – among the three categories in current space-communication studies. For the two types of the studies, we unearthed spatial properties rarely utilized so far and proposed new metrics for them. Although the analysis at the layout-level is one of the core themes in this field, most studies at the layout-level have focused on the *openness* of the layout, and the openness is usually operationalized with a categorical variable classifying a layout as ‘open’ or ‘closed’. Thus, the properties of a layout other than openness have been often neglected. Moreover, even openness has not been properly addressed because such a categorical variable hides subtle differences in the degree of openness within a category.

A few studies (Sailer 2010; Toker and Gray 2008; Peponis et al. 2007) conducted layout-level analysis without such a categorical variable for openness. They quantitatively described their layouts mainly by taking an average of a metric for a space such as ‘integration’ (average distance to other spaces from the focal space) or ‘connectivity’ (the

number of connected spaces to the focal space). It is true that such metrics provide better quantified descriptions of a layout than categorical variables. However, such layout-level metrics do not consider the locations of building users and key facilities, which is found to be invaluable in the analysis of the effect of spatial layout.

Further research should therefore concentrate on the investigation of methods for layout-level analysis. First, we need to theorize the relevant properties of workplace layout to communications. Then we need to develop 'native' metrics for the properties, not derived from the metrics developed for a space-level analysis. Finally, we should demonstrate the validity of the developed metrics with empirical evidence. Metrics measuring the property of a layout would be particularly assistive to designers because such metrics allow designers to easily compare their design alternatives.

***Shortest path choice vs Probabilistic path choice***

It would be interesting to compare the 'shortest path choice model' and the 'probabilistic path choice model'. We considered only the shortest paths in calculating sociospatial betweenness and zone overlap. In calculating sociospatial betweenness, for example, a social tie between a dyad is spatialized along with the shortest path in the spatial network. In calculating zone overlap, the functional zone of each person consists of key facilities and the shortest paths between the key facilities. To use the shortest paths in such ways implicitly assumes that people's movements always follow the shortest path in the spatial network. This assumption makes the model simpler, but it might be too strong an assumption. To relax the assumption, we may adopt the 'probabilistic choice model' where shorter paths are more likely to be chosen while longer paths are less likely to be chosen as suggested in random walk betweenness (Newman 2005).

***Extension of functional zone and the SPSN***

The dissertation introduces two unique spatial concepts upon which the two metrics are developed: the *functional zone* for zone overlap, and the *spatial projection of a social network* (SPSN) for sociospatial betweenness. As we discussed in each chapter, the use of the two concepts are not limited to the construction of the two metrics.

The concept of the functional zone effectively illustrates a person's individual sphere of operation in the workplace. If all people's functional zones are overlaid on the floor and the number of total overlaps in each space is counted, the number can be a proxy of how many diverse people use the space regularly. It would be interesting to investigate how the number of overall overlaps of a space compare to the patterns of space utilization of the space.

The possibility of the SPSN as a broader platform of sociospatial analysis would be worth exploring. Although we developed and investigated only sociospatial betweenness in this study, we can derive other sociospatial centralities from the SPSN in a similar manner as we derived sociospatial betweenness. We briefly outlined how other sociospatial centralities such as sociospatial closeness and sociospatial degree can be defined using the SPSN in Chapter 3. Further investigation on these sociospatial centralities would be an interesting topic.

***Development of Integrated software***

Another possible area of future research would be to develop integrated software for analyzing spatial properties at the all three levels of the analysis. Although there are several software packages for spatial network analysis, most of the current packages provide tools only for space-level analyses. Moreover, the current packages have limited

functionality for managing non-spatial information such as information on individual occupants, the social relationship among them, and the organizational structure.

Thus further effort should be made to develop a software package calculating metrics for all three levels: space-level, dyad-level, layout-level analysis including the new metrics proposed in this study. Also the package would be favorable if it can manage and visualize not only spatial data, but also user data such as the location of users, survey responses, and administrative data; information on space utilization such as observed use of each space and the locations of key facilities; lastly, data on social relationships such as organizational hierarchy and friendship ties.

We may develop such a software package by extending an existing GIS framework so that it can manage various types of information within the package. Adopting a GIS framework will also provide additional advantages such as functions for general spatial operation and querying, customizable data visualization, compatibility with widely used file formats, and familiarity to existing users.

# BIBLIOGRAPHY

# Bibliography


Agneessens, Filip, and Rafael Wittek. 2012. “Where Do Intra-Organizational Advice Relations Come from? The Role of Informal Status and Social Capital in Social Exchange.” *Social Networks* 34 (3): 333–45.

Allen, Thomas J. 1977. *Managing the Flow of Technology: Technology Transfer and the Dissemination of Technological Information within the R.& D.Organization*. MIT Press.

Allen, Thomas J., and A. R. Fustfeld. 1975. “Research Laboratory Architecture and the Structuring of Communications.” *R&D Management* 5 (2): 153–64. doi:10.1111/j.1467-9310.1975.tb01230.x.

Allen, Thomas J., and Peter G. Gerstberger. 1973. “A Field Experiment to Improve Communications in a Product Engineering Department: The Nonterritorial Office.” *Human Factors: The Journal of the Human Factors and Ergonomics Society* 15 (5): 487–98. doi:10.1177/001872087301500505.

Appel-Meulenbroek, Rianne. 2010. “Knowledge Sharing through Co-Presence: Added Value of Facilities.” *Facilities* 28 (3/4): 189–205. doi:10.1108/02632771011023140.

Boje, Axel. 1971. *Open-Plan Offices*. Random House Business.

Boutellier, Roman, Fredrik Ullman, Jürg Schreiber, and Reto Naef. 2008. “Impact of Office Layout on Communication in a Science-Driven Business.” *R&D Management* 38 (4): 372–91. doi:10.1111/j.1467-9310.2008.00524.x.

Brandes, Ulrik. 2008. “On Variants of Shortest-Path Betweenness Centrality and Their Generic Computation.” *Social Networks* 30 (2): 136–45. doi:10.1016/j.socnet.2007.11.001.

Brookes, Malcolm J., and Archie Kaplan. 1972. “The Office Environment: Space Planning and Affective Behavior.” *Human Factors: The Journal of the Human Factors and Ergonomics Society* 14 (5): 373–91. doi:10.1177/001872087201400502.

Brown, Eliot. 2012. “Gehry Designs Facebook Office Fit for ‘Buzz’ — In One Room.” *The Wall Street Journal*, August 29. http://blogs.wsj.com/developments/2012/08/29/gehry-designs-facebook-office-fit-for-buzz-in-one-room/.

Burt, R. S. 2004. “Structural Holes and Good Ideas.” *AJS* 110 (2): 349–99.

Butts, Carter T., Ryan M. Acton, John R. Hipp, and Nicholas N. Nagle. 2012. “Geographical Variability and Network Structure.” *Social Networks* 34 (1): 82–100. doi:10.1016/j.socnet.2011.08.003.

Chen, Jingchun, and Bo Yuan. 2006. “Detecting Functional Modules in the Yeast Protein–protein Interaction Network.” *Bioinformatics* 22 (18): 2283–90. doi:10.1093/bioinformatics/btl370.

Clearwater, Y. 1980. “Comparison of Effects of Open and Closed Office Designs on Job Satisfaction and Productivity.”

Cowgill, Bo, Justin Wolfers, and Eric Zitzewitz. 2009. “Using Prediction Markets to Track Information Flows: Evidence from Google.” In *AMMA*, 3. http://www.msenergysolutions.net/index.pl/ar/00/http/www.bocowgill.com/GooglePredictionMarketPaper.pdf.

Daraganova, Galina, Pip Pattison, Johan Koskinen, Bill Mitchell, Anthea Bill, Martin Watts, and Scott Baum. 2012. “Networks and Geography: Modelling Community Network Structures as the Outcome of Both Spatial and Network Processes.” *Social Networks* 34 (1): 6–17. doi:10.1016/j.socnet.2010.12.001.

Dijkstra, E. W. 1959. “A Note on Two Problems in Connexion with Graphs.” *Numerische Mathematik* 1 (1): 269–71.

Ehrenberg, Ronald G., and Liang Zhang. 2005. “Do Tenured and Tenure-Track Faculty Matter?” *Journal of Human Resources* 40 (3): 647–59.

Elsbach, Kimberly D., and Michael G. Pratt. 2007. “4 The Physical Environment in Organizations.” *The Academy of Management Annals* 1 (1): 181–224. doi:10.1080/078559809.

Festinger, Leon, Stanley Schachter, and Kurt Back. 1950. *Social Pressures in Informal Groups: A Study of Human Factors in Housing*. Stanford Univ Pr.

Fortunato, Santo. 2010. “Community Detection in Graphs.” *Physics Reports* 486 (3-5): 75–174. doi:10.1016/j.physrep.2009.11.002.

Goldberger, Paul. 2013. “Google’s New Built-from-Scratch Googleplex.” *Vanity Fair*, February 22. http://www.vanityfair.com/online/daily/2013/02/exclusive-preview-googleplex.

Grannis, Rick. 2009. *From the Ground Up: Translating Geography into Community through Neighbor Networks*. Princeton University Press.

Gullahorn, John T. 1952. “Distance and Friendship as Factors in the Gross Interaction Matrix.” *Sociometry* 15 (1/2): 123–34. doi:10.2307/2785450.

Hansen, M. T. 1999. “The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits.” *Administrative Science Quarterly* 44 (1): 82–85.

Hanson, Lady Alice. 1983. “Effects of a Move to an Open Landscape Office”. Claremont Graduate School.

Hatch, Mary Jo. 1987. “Physical Barriers, Task Characteristics, and Interaction Activity in Research and Development Firms.” *Administrative Science Quarterly* 32 (3): 387–99.

Heerwagen, Judith H, Kevin Kampschroer, Kevin M Powell, Vivian Loftness, and Charles M Salter. 2004. “Collaborative Knowledge Work Environments.” *Building Research & Information* 32 (6): 512–28.

Hillier, Bill. 1996. *Space Is the Machine: A Configurational Theory of Architecture*. Cambridge: Cambridge University Press.

Hillier, Bill, and Julienne Hanson. 1984. *The Social Logic of Space*. Reprint. Cambridge University Press.

Hillier, Bill, and Alan Penn. 1991. “Visible Colleges: Structure and Randomness in the Place of Discovery.” *Science in Context* 4 (01): 23–50. doi:10.1017/S0269889700000144.

Hipp, John R., Robert W. Faris, and Adam Boessen. 2012. “Measuring ‘neighborhood’: Constructing Network Neighborhoods.” *Social Networks* 34 (1): 128–40. doi:10.1016/j.socnet.2011.05.002.

Hua, Ying, Vivian Loftness, Judith H. Heerwagen, and Kevin M. Powell. 2010. “Relationship Between Workplace Spatial Settings and Occupant-Perceived Support for Collaboration.” *Environment and Behavior*, October. doi:10.1177/0013916510364465.

Jackson, Matthew O. 2008. *Social and Economic Networks*. Princeton University Press.

Kahn, Arnold, and Timothy A. McGaughey. 1977. “Distance and Liking: When Moving Close Produces Increased Liking.” *Sociometry* 40 (2): 138–44. doi:10.2307/3033517.

Karazsia, Bryan T., and Manfred HM van Dulmen. 2008. “Regression Models for Count Data: Illustrations Using Longitudinal Predictors of Childhood Injury.” *Journal of Pediatric Psychology* 33 (10): 1076–84.

Karp, Richard M. 1972. “Reducibility among Combinatorial Problems.” In *Complexity of Computer Computations*, edited by Raymond E. Miller, James W. Thatcher, and

Jean D. Bohlinger, 85–103. The IBM Research Symposia Series. Springer US. http://link.springer.com/chapter/10.1007/978-1-4684-2001-2_9.

Kembel, Steven W., James F. Meadow, Timothy K. O'Connor, Gwynne Mhuireach, Dale Northcutt, Jeff Kline, Maxwell Moriyama, G. Z. Brown, Brendan J. M. Bohannan, and Jessica L. Green. 2014. "Architectural Design Drives the Biogeography of Indoor Bacterial Communities." *PLoS ONE* 9 (1): e87093. doi:10.1371/journal.pone.0087093.

Kono, Clifford, Donald Palmer, Roger Friedland, and Matthew Zafonte. 1998. "Lost in Space: The Geography of Corporate Interlocking Directorates 1." *American Journal of Sociology* 103 (4): 863–911.

Kossinets, Gueorgi, and Duncan J. Watts. 2006. "Empirical Analysis of an Evolving Social Network." *Science* 311 (5757): 88–90.

Kraut, Robert, Carmen Egido, and Jolene Galegher. 1988. "Patterns of Contact and Communication in Scientific Research Collaboration." In *Proceedings of the 1988 ACM Conference on Computer-Supported Cooperative Work*, 1–12. CSCW '88. New York, NY, USA: ACM. doi:10.1145/62266.62267.

Lee, Kyungjoon, John S. Brownstein, Richard G. Mills, and Isaac S. Kohane. 2010. "Does Collocation Inform the Impact of Collaboration?" *PLoS ONE* 5 (12): e14279. doi:10.1371/journal.pone.0014279.

Lehrer, Jonah. 2012. *Imagine: How Creativity Works*. 1st edition. Houghton Mifflin.

Li, Xiang, Christophe Claramunt, and Cyril Ray. 2010. "A Grid Graph-Based Model for the Analysis of 2D Indoor Spaces." *Computers, Environment and Urban Systems*, GeoVisualization and the Digital City Special issue of the International Cartographic Association Commission on GeoVisualization, 34 (6): 532–40. doi:10.1016/j.compenvurbsys.2010.07.006.

Liu, Christopher C. 2010. "A SPATIAL ECOLOGY OF STRUCTURAL HOLES: SCIENTISTS AND COMMUNICATION AT A BIOTECHNOLOGY FIRM." In

*Academy of Management Proceedings*, 2010:1–6. Academy of Management. http://proceedings.aom.org/content/2010/1/1.175.short.

Long, J. Scott, and Jeremy Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata, Second Edition*. Stata Press.

Miller, Claire Cain, and Catherine Rampell. 2013. "Yahoo Orders Home Workers Back to the Office." *The New York Times*, February 25, sec. Technology. http://www.nytimes.com/2013/02/26/technology/yahoo-orders-home-workers-back-to-the-office.html.

Mizruchi, Mark S., and Christopher Marquis. 2006. "Egocentric, Sociocentric, or Dyadic?: Identifying the Appropriate Level of Analysis in the Study of Organizational Networks." *Social Networks* 28 (3): 187–208.

Monge, Peter R., and Kenneth K. Kirste. 1980. "Measuring Proximity in Human Organization." *Social Psychology Quarterly* 43 (1): 110–15. doi:10.2307/3033753.

Moodysson, Jerker, and Ola Jonsson. 2007. "Knowledge Collaboration and Proximity the Spatial Organization of Biotech Innovation Projects." *European Urban and Regional Studies* 14 (2): 115–31.

Newman, M. E. J. 2006. "Modularity and Community Structure in Networks." *Proceedings of the National Academy of Sciences* 103 (23): 8577–82.

Newman, M. E. J., and M. Girvan. 2004. "Finding and Evaluating Community Structure in Networks." *Physical Review E* 69 (2): 026113. doi:10.1103/PhysRevE.69.026113.

Newman, M.E. J. 2005. "A Measure of Betweenness Centrality Based on Random Walks." *Social Networks* 27 (1): 39–54. doi:10.1016/j.socnet.2004.11.009.

Obstfeld, D. 2005. "Social Networks, the Tertius Iungens Orientation, and Involvement in Innovation." *Administrative Science Quarterly* 50 (1): 100–130.

Oldham, Greg R., and Daniel J. Brass. 1979. “Employee Reactions to an Open-Plan Office: A Naturally Occurring Quasi-Experiment.” *Administrative Science Quarterly* 24 (2): 267–84. doi:10.2307/2392497.

Penn, Alan, J. Desyllas, and L. Vaughan. 1999. “The Space of Innovation: Interaction and Communication in the Work Environment.” *Environment and Planning B: Planning and Design* 26 (2): 193–218. doi:10.1068/b260193.

Peponis, John. 1985. “The Spatial Culture of Factories.” *Human Relations* 38 (4): 357 – 390. doi:10.1177/001872678503800405.

Peponis, John, S. Bafna, R. Bajaj, J. Bromberg, C. Congdon, Mahbub Rashid, S. Warmels, Yan Zhang, and Craig Zimring. 2007. “Designing Space to Support Knowledge Work.” *Environment and Behavior* 39 (6): 815–40. doi:10.1177/0013916506297216.

Peponis, John, and Jean D. Wineman. 2002. “Spatial Structure of Environment and Behavior.” In *Handbook of Environmental Psychology*, 271–91.

Peponis, John, Jean D. Wineman, M. Rashid, S. Kim, and S. Bafna. 1997. “On the Description of Shape and Spatial Configuration inside Buildings: Convex Partitions and Their Local Properties.” *Environment and Planning B* 24: 761–82.

Pfeffer, Jeffrey. 1992. *Managing with Power: Politics and Influence in Organizations*. Harvard Business Press.

Pierce, Charles A., Donn Byrne, and Herman Aguinis. 1996. “Attraction in Organizations: A Model of Workplace Romance.” *Journal of Organizational Behavior* 17 (1): 5–32.

Porter, Mason A., Jukka-Pekka Onnela, and Peter J. Mucha. 2009. “Communities in Networks.” *Notices of the AMS* 56 (9): 1082–97.

Powell, Gary N., and Sharon Foley. 1998. “Something to Talk About: Romantic Relationships M Organizational Settings.” *Journal of Management* 24 (3): 421–48.

Powell, Walter W., Kenneth W. Koput, and Laurel Smith-Doerr. 1996. “Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology.” *Administrative Science Quarterly* 41 (1): 116–45.

Priest, Robert F., and Jack Sawyer. 1967. “Proximity and Peership: Bases of Balance in Interpersonal Attraction.” *American Journal of Sociology* 72 (6): 633–49.

Rashid, Mahbub, Kevin Kampschroer, Jean D. Wineman, and Craig Zimring. 2006. “Spatial Layout and Face-to-Face Interaction in Offices—a Study of the Mechanisms of Spatial Effects on Face-to-Face Interaction.” *Environment and Planning B: Planning and Design* 33 (6): 825–44. doi:10.1068/b31123.

Rosvall, Martin, and Carl T. Bergstrom. 2008. “Maps of Random Walks on Complex Networks Reveal Community Structure.” *Proceedings of the National Academy of Sciences* 105 (4): 1118–23. doi:10.1073/pnas.0706851105.

Sack, Robert David. 1986. *Human Territoriality: Its Theory and History*. CUP Archive.

———. 1993. “The Power of Place and Space.” *Geographical Review*, 326–29.

Sailer, Kerstin. 2010. “The Space-Organisation Relationship: On the Shape of the Relationship between Spatial Configuration and Collective Organisational Behaviours”. TU Dresden.

Sailer, Kerstin, Andrew Budgen, Nathan Lonsdale, Alasdair Turner, and Alan Penn. 2009. “Comparative Studies of Offices Pre and Post How Changing Spatial Configurations Affect Organisational Behaviours.” In *Proceedings of the 7th International Space Syntax Symposium*, 1–14. Stockholm.

Sailer, Kerstin, and Ian McCulloh. 2012. “Social Networks and Spatial configuration—How Office Layouts Drive Social Interaction.” *Social Networks* 34 (1): 47–58. doi:10.1016/j.socnet.2011.05.005.

Silverman, Rachel Emma. 2013. “The Science of Serendipity in the Workplace.” *Wall Street Journal*, May 1, sec. Management. http://online.wsj.com/news/articles/SB10001424127887323798104578455081218505870.

Snyder, Thomas D., and Sally A. Dillow. 2012. *Digest of Education Statistics 2011*. Government Printing Office.

Steen, Jesper, and Henrik Markhede. 2010. “Spatial and Social Configurations in Offices.” *The Journal of Space Syntax* 1 (1): 121–32.

Stryker, J.B., M.D. Santoro, and G.F. Farris. 2012. “Creating Collaboration Opportunity: Designing the Physical Workplace to Promote High-Tech Team Communication.” *IEEE Transactions on Engineering Management* 59 (4): 609–20. doi:10.1109/TEM.2011.2170995.

Sundstrom, Eric, R. Kring Herbert, and David W. Brown. 1982. “Privacy and Communication in an Open-Plan Office A Case Study.” *Environment and Behavior* 14 (3): 379–92. doi:10.1177/0013916582143007.

Sykes, Richard E. 1977. *A Theory of Proximity and Attraction*. Minnesota Systems Research.

Toker, U., and D. O Gray. 2008. “Innovation Spaces: Workspace Planning and Innovation in US University Research Centers.” *Research Policy* 37 (2): 309–29.

Turner, A., and Alan Penn. 1999. “Making Isovists Syntactic: Isovist Integration Analysis.” In *2nd International Symposium on Space Syntax, Brasilia*.

Tyler, Joshua R., Dennis M. Wilkinson, and Bernardo A. Huberman. 2005. “E-Mail as Spectroscopy: Automated Discovery of Community Structure within

Organizations." *The Information Society* 21 (2): 143–53. doi:10.1080/01972240590925348.

Van Eck, Nees Jan, and Ludo Waltman. 2007. "VOS: A New Method for Visualizing Similarities Between Objects." In *Advances in Data Analysis*, edited by Professor Dr Reinhold Decker and Professor Dr Hans-J. Lenz, 299–306. Studies in Classification, Data Analysis, and Knowledge Organization. Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-540-70981-7_34.

Waltman, Ludo, Nees Jan van Eck, and Ed C.M. Noyons. 2010. "A Unified Approach to Mapping and Clustering of Bibliometric Networks." *Journal of Informetrics* 4 (4): 629–35. doi:10.1016/j.joi.2010.07.002.

Wineman, Jean D., Yongha Hwang, Felichism Kabo, Jason Owen-Smith, and Gerald F. Davis. 2014. "Spatial Layout, Social Structure, and Innovation in Organizations." *Environment and Planning B: Planning and Design* 41: 000–000.

Wineman, Jean D., F. W Kabo, and G. F Davis. 2008. "Spatial and Social Networks in Organizational Innovation." *Environment and Behavior* 41 (3): 427.

Zajonc, Robert B. 1968. "Attitudinal Effects of Mere Exposure." *Journal of Personality and Social Psychology* 9 (2, Pt.2): 1–27. doi:10.1037/h0025848.