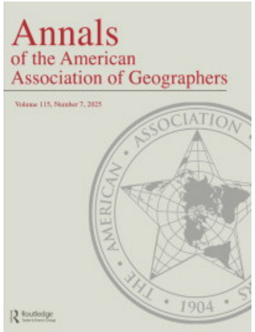

Taylor & Francis
Taylor & Francis Group

Annals of the American Association of Geographers

ISSN: 2469-4452 (Print) 2469-4460 (Online) Journal homepage: www.tandfonline.com/journals/raag21

Modeling Region Affiliation with Fuzzy Membership Based on Spatial and Social Interactions

Jacob Kruse, Song Gao & Kenneth R. Mayer

To cite this article: Jacob Kruse, Song Gao & Kenneth R. Mayer (10 Sep 2025): Modeling Region Affiliation with Fuzzy Membership Based on Spatial and Social Interactions, Annals of the American Association of Geographers, DOI: 10.1080/24694452.2025.2551044

To link to this article: https://doi.org/10.1080/24694452.2025.2551044

Published online: 10 Sep 2025.

ANNALS OF THE AMERICAN ASSOCIATION OF GEOGRAPHERS
https://doi.org/10.1080/24694452.2025.2551044

Modeling Region Affiliation with Fuzzy Membership Based on Spatial and Social Interactions

Jacob Kruse,[a] Song Gao,[a] and Kenneth R. Mayer[b]

[a]*Geospatial Data Science Lab, Department of Geography, University of Wisconsin, Madison, USA;* [b]*Department of Political Science, University of Wisconsin, Madison, USA*

A key challenge in regionalization is that regions, such as urban function zones or climate zones, often have indeterminate boundaries, making it difficult to exactly quantify their geographic extent. Political redistricting, as a regionalization task, deals with this problem acutely, as requirements to preserve communities of interest (COIs) do not define such communities, introducing inherent vagueness in their boundaries. To address this issue, this work introduces a network approach that models COIs by integrating spatial-social interactions and evaluates district assignment by quantifying the degree to which a geographic area is connected to all other areas within each district. Furthermore, we draw on a splatial framework to understand the different spaces in which modern human communities interact, allowing us to more comprehensively model the community interactions that constitute COIs by using both spatial and social interactions, as measured with human mobility flows and social network connections. By comparing how district membership aligns across these two interaction types with the fuzzy membership methodology, it can reveal distinct spatial patterns, while combining them can reduce ambiguity in region membership. To demonstrate its utility, the proposed methodology is applied to a 2020 congressional district plan for the State of Wisconsin. Beyond redistricting, this work also contributes to the geography literature by providing a spatial interaction-based framework for quantifying regional affiliations in boundary areas. *Key Words: fuzzy membership, redistricting, regionalization, spatial networks, spatial-social interactions.*

The delineation of regional boundaries is a foundational issue in geography, and a key challenge in this process is that the objects we seek to model as regions, such as places (Montello et al. 2003; Montello, Friedman, and Phillips 2014; Gao et al. 2017; Bae and Montello 2018) or landscapes (Brown 1998; Hall and Arnberg 2002), often have indeterminate boundaries (Burrough and Frank 1996; Leung 1999). This indeterminacy arises from spatial, semantic, or ontological (Varzi 2001; Smith and Mark 2003) vagueness in the definitions of such objects and their spatial relationships (Q. Guo, Liu, and Wieczorek 2008; Ding et al. 2025). Political redistricting, as a regionalization task, deals with this problem acutely, as political boundaries are expected to preserve communities of interest (COIs)—groups with shared concerns that are likely to be affected by legislation (Malone 1997)—even though such communities are not clearly defined by states (Forest 2004; Webster 2013). Various definitions have been proposed in the academic literature and used in legal challenges to existing district boundaries, but they all suffer from semantic (S. J. Chen et al. 2022) and ontological (Stephanopoulos 2012a) vagueness. This vagueness in defining the boundaries of the underlying communities inherently creates ambiguity regarding the appropriate placement of district boundaries.

To address the issue of vagueness in regionalization, the GIScience literature has increasingly turned to social sensing methods, which integrate multiple data sources—such as social media and human mobility data—to characterize places (Gao et al. 2017; Zhu et al. 2020) and activity spaces (Jin et al. 2021; Z. Chen, Zou, and Tan 2025), and define regions with greater clarity and consistency (Liu et al. 2015; Liao et al. 2018). These approaches generate composite membership scores for geographic

ARTICLE HISTORY
Initial submission, January 2025; revised submissions, June and July 2025; final acceptance, July 2025

CORRESPONDING AUTHOR Song Gao song.gao@wisc.edu

subunits, helping to distinguish core areas of strong agreement from peripheral or contested zones (McKenzie and Adams 2017; W. Chen et al. 2018).

Another membership-based approach to addressing region ambiguity is fuzzy logic (Zadeh 1965), which has been used extensively for modeling spatial objects with indeterminate boundaries in GIScience (Cohn and Gotts 1996; Fonte and Lodwick 2004; Shi and Liu 2004). Rather than assigning each area exclusively to a single region, the fuzzy membership approach calculates the partial membership that each area has in all regions, after which a membership cutoff value is applied to determine final region assignment (Burrough and McDonnell 1989; Ahmed et al. 2020). Although various fuzzy approaches to region delineation have been described in the literature (Cohn and Gotts 1996; Erwig and Schneider 1997; Liu, Yuan, and Gao 2019), these approaches predominantly rely on attribute values at specific locations. There is currently no fuzzy approach, however, that addresses vagueness in regionalization when the region memberships are defined with spatial interactions between areas.

Spatial interactions are movements and information exchanges across space driven by human processes (Ratti et al. 2010; Nelson and Rae 2016; Kang et al. 2020; Xu et al. 2022). Because they effectively capture human interests, geographers have used spatial interactions in numerous regionalization tasks (Holtz et al. 2020; M. Li et al. 2021; Liang et al. 2022) to create regions that better reflect underlying human dynamics and contribute to the understanding of geographic process-oriented regionalization (Zhang et al. 2024). Despite these advancements, though, the application of spatial interaction networks in political redistricting to understand COIs remains relatively underexplored. Two recent works in particular (Kruse, Gao, Ji, Levin, et al. 2024; Kruse, Gao, Ji, Szabo, et al. 2024) have used spatial interaction networks, derived from human mobility flows, to model COIs in redistricting. In these studies, spatial interactions represent the shared economic and civic life that state courts generally understand as constituting a COI (Stephanopoulos 2012a). Existing methods, however, do not address the common need to evaluate whether a given area is placed in a district with other areas to which it is strongly connected—in other words, whether COIs are preserved by district boundaries (S. J. Chen et al. 2022).

Additionally, while existing works have used different types of spatial interactions to identify regions, a conceptual framework for the use of spatial interactions to represent the various human activities that constitute communities is lacking. For such a framework, we turn to the splatial framework developed in Shaw and Sui (2020), which offers a valuable conceptual lens for analyzing contemporary human dynamics. By incorporating multiple dimensions of space—including absolute, relative, relational, and mental space—this framework helps contextualize the different types of interactions that define modern human dynamics, such as physical movements in absolute space and social connections in relational space.

In this work, we present a network-based fuzzy membership approach to modeling COIs in redistricting, which, more broadly, can be used to model regions based on community interactions. The main contributions are threefold. First, we employ the splatial framework (Shaw and Sui 2020) to justify why multitype spatial interactions can better model communities, demonstrating the robustness of this approach using spatial and social interactions. Second, to operationalize this new definition of COIs for the purposes of redistricting, we present a fuzzy membership approach that quantifies region affiliation based on the multitype interactions. In this approach, the strength of interactions, or *affiliation*, between a given geographic area (e.g., a census block group [CBG]) and a region (e.g., a political district composed of multiple CBGs) can be assessed, enabling an evaluation of whether the area is assigned to the region with which it has the strongest connections. Third, we show how the fuzzy membership approach can be used to quantify the extent to which membership ambiguity is reduced through the combination of two types of spatial interactions. Overall, this work benefits research on COIs in redistricting, as well as regionalization with spatial interaction communities, more broadly.

The remainder of the article is organized as follows. We begin by reviewing the relevant literature on regionalization and COIs, multitype spatial interactions, the splatial conceptualization of space, social sensing, and fuzzy membership. We then introduce the new methodology for measuring region affiliation and district plan evaluation. The results show how the proposed methodology can be applied in the detailed analysis of a particular congressional redistricting plan, and how it can be applied to compare one plan to a distribution of valid alternatives. We then address some of the limitations of this work and provide concluding remarks.

Related Work

Regionalization and COIs

Regionalization is the process of partitioning geographic space into spatially coherent and contiguous units, called regions, that are internally homogeneous according to specified criteria. These criteria can include demographic, socioeconomic, ecological, health, or interaction-based measures, among others (James 1952; Storper 1997; Helbich et al. 2013; Mu et al. 2014; Bae and Montello 2018; Liang et al. 2022). Unlike nonspatial clustering approaches, regionalization explicitly enforces geographic constraints—such as contiguity and compactness—to ensure regions form spatially continuous and geographically interpretable units (D. Guo 2008; DeFord, Duchin, and Solomon 2019; Aydin et al. 2021; Liang et al. 2025). This spatial continuity aligns with the ontology of regions as coherent and meaningful geographic entities, where adjacency and proximity play crucial roles in their definition. By imposing these constraints, regionalization methods yield clusters that better reflect real-world geographic structures, enhancing their interpretability and practical utility for applications like political redistricting, resource management, and urban planning (Brown 1998; Hall and Arnberg 2002; Montello et al. 2003; Duque, Anselin, and Rey 2012).

Political redistricting, as a specific type of regionalization, faces the additional challenge of preserving COIs—that is, ensuring that such communities are not divided by political district boundaries (Malone 1997). This task is further complicated by the fact that COIs are often not clearly defined, creating significant practical difficulties (Forest 2004; Webster 2013). Borrowing from the broader regionalization literature, these challenges can be explicitly described using the concept of *vagueness*, which includes two distinct phenomena: *semantic vagueness*, relating to ambiguity in terminology (e.g., what precisely does "downtown" mean), and *ontological vagueness*, arising from unclear or gradual boundaries in the physical or social phenomena being partitioned (e.g., where exactly does a forest end; Leung 1999; Bennett 2001; Varzi 2001; Montello et al. 2003; Smith and Mark 2003; Gao et al. 2013). Thus, the lack of clarity in defining COIs amplifies the inherent semantic and ontological challenges of political redistricting.

Regarding semantic ambiguity, many states require the preservation of COIs, yet no states clearly define these communities (Forest 2004; Webster 2013). To address this ambiguity, various definitions have been proposed in academic literature or employed in legal arguments. For instance, Stephanopoulos (2012a) argued that legal precedent supports the concept of a territorial community, defining a COI as a spatially bounded region characterized by relatively homogeneous sociodemographic traits, culture, and industry. Although this definition emphasizes the geographic aspect of COIs, it remains unclear exactly how similar a given area must be a broader COI to justify its inclusion. Similarly, metrics such as spatial diversity, used to quantify district homogeneity (Stephanopoulos 2012b), do not account for how homogeneity changes over the spatial extent of a district—a critical consideration, given that natural and social phenomena frequently change continuously over space (Burrough and Frank 1996; Cohn and Gotts 1996).

In another approach to modeling COIs, constituents are invited to self-identify COI boundaries using geographic information systems (GIS) software (Makse 2012; S. J. Chen et al. 2022). Although this method can reflect the groupings that constituents consider meaningful, it is heavily influenced by the demographics and participation of the individuals contributing to the identification process. Additionally, it is limited by the semantic vagueness of the term COI itself, which lacks a universally agreed-on definition. Other approaches to defining COIs have emerged in the context of legal challenges to district boundaries. In particular, litigants in several cases have relied on social and economic linkages, along with geographic and historical reasoning, to argue that COIs were either unnaturally split or improperly grouped together (S. J. Chen et al. 2022). The lack of a clear, standardized framework for analyzing the linkages that define COIs, however, complicates the objective evaluation of such claims. Following Tobler's First Law of Geography—that near things are more similar than distant things (Tobler 1970)—it can be argued that community connections and similarities are better represented on a continuum of degree rather than as binary states of existence or nonexistence. Consequently, modeling COIs in a spatially continuous manner could offer a more objective and nuanced approach.

Similar to legal arguments that advocate for the existence of COIs based on social and economic linkages, recent works use spatial interactions to quantify COIs with the understanding that connections between areas can help define the degree to which two areas belong to the same COI. As described earlier, spatial interactions are movements or exchanges across space driven by human processes. These interactions encompass a diverse set of phenomena, including human mobility flows captured through surveys (Nelson and Rae 2016), mobile phone calls, or location-tracking data (Ratti et al. 2010; Kang et al. 2020; Xu et al. 2022), as well as the flows of commodities, information, and social connections (B. Li et al. 2020; Rao et al. 2022). In studies using spatial interactions for regionalization, the concept of interaction communities is employed, where higher volumes of spatial interactions between regions serve as indicators of stronger social and economic cohesion (Gao et al. 2013; Dong et al. 2015; Y. Chen, Zhang, and Liang 2019).

To the best of our knowledge, only two works in the literature have applied spatial interactions to redistricting. In one study, Kruse, Gao, Ji, Szabo, et al. (2024) employed the interaction ratio to provide a plan-wide assessment of how spatial interactions are directed within versus between districts. This approach enables comparisons between district maps to evaluate how well boundaries align with underlying spatial interaction communities. Although this method is effective for comparing plans, it does not facilitate the analysis of a single redistricting map at the subdistrict level, such as evaluating whether geographically neighboring areas are appropriately grouped within the same district. In another study, Kruse, Gao, Ji, Levin, et al. (2024) evaluated the strength and consistency of spatial interactions between members of key subnetworks within districts over time. Although this work examines connections within specific subdistrict areas, it does not compare these internal connections to those spanning district boundaries, limiting its utility for determining whether district plans inappropriately split coherent COIs. Notably, both of these works use human mobility flows as the spatial interaction studied.

Conceptualizing Spatial Interactions through the Splatial Framework

Although the previously mentioned studies (Kruse, Gao, Ji, Levin, et al. 2024; Kruse, Gao, Ji, Szabo, et al. 2024) employ spatial interactions to model COIs, they do not provide a framework for understanding how different spatial interaction types can be used to model human communities. For such a framework, we employ the splatial framework developed by Shaw and Sui (2020). In that work, the authors noted how spatial interactions (which they referred to as human dynamics) have rapidly changed with the societal-wide penetration of smart technologies (e.g., smartphones), which have drastically changed how people interact, creating a new world in which "physical and virtual, objective and subjective, territorial and topological worlds are increasingly coupled and entangled for most human activities" (442). The authors synthesized various conceptualizations of both space and place to describe human interactions as occurring in a variety of overlapping spaces and places. This space–place (splatial) framework includes absolute, relative, relational, and mental spaces.

Using this framework, the previous works on redistricting can be considered as using interactions that occur in absolute space. Kruse, Gao, Ji, Levin, et al. (2024) and Kruse, Gao, Li, Szabo, et al. (2024) used human mobility flows between CBGs, and are thus interactions that occur in absolute space. Human mobility flows could be considered as occurring in relational space, but we here consider them as occurring in absolute space, as the strength of flows is strongly related to the geographic distance between origin and destination, a result of the cost of travel in terms of time and money.

Given that spatial interactions (i.e., human activities and interactions) can occur in multiple types of spaces in the modern era, redistricting works employing interactions in multiple types of spaces have the potential to create more robustly defined interaction communities, as they encompass a wider range of human activities in their measurement. In the age of smart technology, one salient space of human activity is virtual space (Shaw and Yu 2009), particularly regarding online social networks such as Facebook (Meta). In the splatial framework, such social networks are considered as occurring in relational space, as topology, rather than absolute coordinates, is what defines the space. We also note that the attempts at modeling COIs wherein constituents self-identify COI boundaries using GIS software (Makse 2012; S. J. Chen et al. 2022) are employing the mental space of the splatial framework. Even if this approach could be combined with spatial interactions to describe COIs more comprehensively, it tends to be heavily biased by whoever is included in the COI survey. Regarding relative space, Shaw and Sui (2020) described it as a movable dimension of absolute space, such as the location of objects relative to an autonomous vehicle. At present, we do not see a use for this space in defining interaction communities in redistricting.

COIs as Regions with Indeterminate Boundaries

As described earlier, current COI definitions face semantic and ontological vagueness stemming from imprecise working definitions and challenges in delineating communities' true geographic expanse (Varzi 2001; Montello et al. 2003; Webster 2013). Although modeling COIs via spatial interactions reduces semantic ambiguity by providing a definition of what constitutes a COI, ontological boundary issues persist. Existing redistricting discourse often assesses if adjacent areas should be in the same district based on cultural homogeneity or strong economic links (Makse 2012; Stephanopoulos 2012b; S. J. Chen et al. 2022). In this case, relatively few pairs are considered. Spatial interactions, however, can lead to $N \times N$ pairwise connections (where N is the number of areas), making boundary indeterminacy a significant, persistent issue due to widespread potential connectivity, as a given area could potentially be connected to all other areas. The use of multitype spatial interactions further complicates the problem. To deal with the continuous nature of these multitype spatial interactions, we turn to social sensing and fuzzy logic methods, which we review next.

Redistricting problems are NP-hard and typically use location-level attributes (Altman 1997; Cannon et al. 2022), unlike the edge attributes of spatial interactions considered here. Thus, we do not aim to create a redistricting algorithm that makes optimal districts. Instead, our research evaluates how well existing district boundaries preserve COIs. Specifically, we assess whether geographic areas are placed in districts where they share most of their spatial-social interactions. This allows for the granular-level assessment of district placement.

Having established the need for methods enabling regionalization based on multiple interaction factors, we now review relevant social sensing and fuzzy membership literature.

Social Sensing

GIScience research has developed social sensing to better define regions and their boundaries (Liu et al. 2015; Cao et al. 2020). Specifically, social sensing combines diverse, human-centered data sources, such as social media interactions and mobile phone records, to generate a more comprehensive representation of regions with ambiguous or overlapping boundaries (Liao et al. 2018). This approach often results in clearer boundaries based on agreement across perspectives (McKenzie and Adams 2017; W. Chen et al. 2018). For instance, Gao et al. (2017) used a data synthesis approach to define vague cognitive regions in California, integrating social media and Internet data sources to generate combined membership scores for all regional subunits. Importantly, this type of approach enables the identification of core versus boundary areas within regions. With social sensing approaches, crisping methods such as the chi-shape are typically used after the membership scores are generated to create clear, binary delineations between regions. In a related work on region discovery, Yuan, Zheng, and Xie (2012) demonstrated how integrating human mobility flows and points of interest can enhance the identification of functional regions, leveraging the distribution of multiple urban functions to provide a more nuanced understanding of each area.

In related areas of study within the field of geography, many works have employed one or multitype spatial interactions to better understand community social structures and regional relationships. For example, Z. Chen, Zou, and Tan (2025) analyzed the nonlinear relationships between migrants' activity-space-based social segregation and traveling distances. Jin et al. (2021) quantified different border effects (natural, infrastructural, and administrative) on intraurban travel through a mobility-based spatial interaction network. Holtz et al. (2020) leveraged both human mobility flows and social media connections to study the impact of uncoordinated COVID-19 responses across regions. By integrating mobility data with social network data, the study quantified how policies in one area influenced geographically and socially connected regions. Similarly, M. Li et al. (2021) proposed a model for predicting human activity intensity that uses graph convolutional networks to incorporate both human mobility flows and cellphone call records between locations. By integrating spatial interaction patterns with social network data, such methods yield more robust results than those obtained using either type of data alone. Although mobility flows and georeferenced social media connections are typically referred to as spatial interactions, we refer to them as *spatial interactions* and *social interactions*, respectively, for clarity throughout the rest of the article.

Fuzzy Membership

Much of the literature on modeling vagueness in GIScience is based on fuzzy logic and fuzzy set theory (Zadeh 1965; Cohn and Gotts 1996; Wang and Hall 1996; Brown 1998), which represent uncertainty by allowing elements to partially belong to multiple classes. To address the aforementioned boundary vagueness challenges in regionalization, GIScience has adopted various fuzzy logic approaches, such as the egg-yolk model, to represent geographic objects or relationships that have inherent vagueness (McBratney and Odeh 1997; Shi and Liu 2004; Liu, Yuan, and Gao 2019). In such models, geographic objects, such as regions, are conceptualized as having core areas of complete membership and transitional boundary zones with partial (but homogeneous) membership (Cohn and Gotts 1996; Wang and Hall 1996; Fonte and Lodwick 2004).

Various regionalization methods have employed fuzzy membership to address the vagueness and uncertainty in determining which region a geographic area should belong to. These methods have been applied in diverse contexts, including the identification of ecosystems and socioeconomic zones (Burrough and McDonnell 1989; Van Ranst et al. 1996; Chavoshi et al. 2013; Ahmed et al. 2020). For instance, Hall and Arnberg (2002) employed fuzzy membership signatures derived from environmental data to delineate landscape regions, allowing the regionalization process to account for gradual transitions between landscape features.

One common fuzzy regionalization method is fuzzy clustering, often implemented with the fuzzy *c*-means algorithm (Hwang and Thill 2009; Goyal and Sharma 2016). This algorithm generalizes traditional *k*-means clustering by assigning each data point a degree of membership to multiple clusters rather than exclusively to one (Bezdek 1981). As a regionalization method, it has been used for a diverse set of problems, including the delineation of housing submarkets (Hwang and Thill 2009), the identification of meteorological drought patterns (Goyal and Sharma 2016), and the analysis of fuzzy soil classes in environmental modeling (McBratney and Odeh 1997). A particularly useful concept in this context is membership entropy, which evaluates the uncertainty or dispersion of membership values in fuzzy clustering (Pal and Bezdek 1995). Higher entropy indicates more evenly distributed membership values across multiple clusters, suggesting ambiguity or uncertainty in the assignment, whereas lower entropy suggests clearer cluster assignments. This approach allows for a quantitative assessment of the membership vagueness of regional subunits. Fuzzy set theory has been specifically applied to redistricting tasks (de Cobos-Silva et al. 2017; Taherdoost and Madanchian 2023), but these existing works primarily focus on aiding the selection of an optimal redistricting plan among multiple options, rather than directly performing regionalization.

In this work we primarily draw on the fuzzy membership literature, but we note that mathematically similar models have been developed in the network science literature, where a probability or strength of membership is assigned to each node for each community. For example, factorization-based methods—such as the BigCLAM affiliation model (Yang and Leskovec 2013)—explicitly represent communities with soft memberships. A comprehensive survey of overlapping community detection algorithms is provided by Xie, Kelley, and Szymanski (2013). More recent work has extended these models to complex settings such as multilayer and attributed networks. For instance, Yang, McAuley, and Leskovec (2013) incorporated node attributes into community detection, whereas Contisciani, De Bacco, and Braunstein (2020) developed a probabilistic model for overlapping community detection in multilayer networks, integrating both network structure and node-level attributes. These approaches demonstrate how allowing soft (or fuzzy) memberships can capture more nuanced community structures in social networks, which conceptually parallels our fuzzy regionalization of spatial-social communities. In geographic studies, researchers have developed novel methods using graph-based deep learning to detect geographically overlapped communities in human mobility spatial networks (Luo and Zhu 2022). To our knowledge, though, none of these overlapping community detection methods incorporate multitype edge weights—such as combining both human mobility flows and social connectedness—as a central input to define membership strength in our research.

Having reviewed the relevant literature, we now outline the methods for incorporating spatial and social interaction networks into the evaluation of political redistricting plans using a fuzzy membership approach.

Methods

Data Sets

To demonstrate how spatial and social interactions can be used to comprehensively understand the regional affiliation of geographic subunits to congressional districts, we analyze the map of Wisconsin 2020 congressional districts produced by the People's Map Commission (PMC), using CBGs as the geographic subunits. The PMC district borders and IDs are shown in Figure 1.

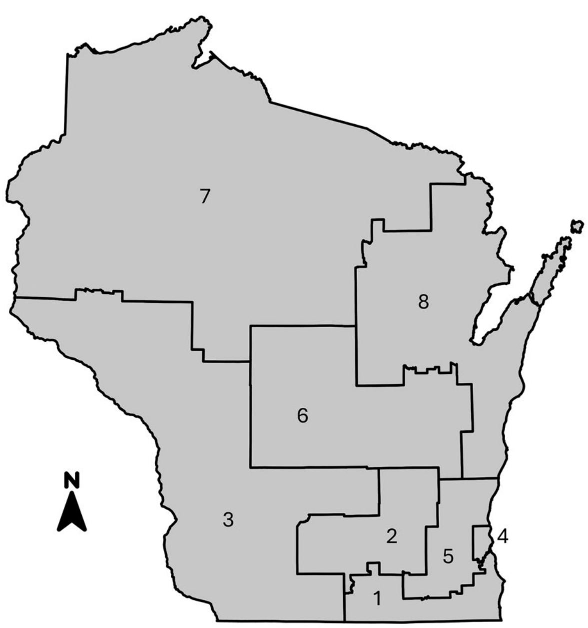

Figure 1. The People's Map Commission congressional district map of Wisconsin.

To measure spatial interactions, we employ the SafeGraph Neighborhood Patterns data set. This data set records the number of anonymized mobile device visits between locations, with visits aggregated at the CBG level (4,489 CBGs in Wisconsin) for each temporal period of the year. Importantly, the recorded trips originate from the users' home locations, ensuring that the data reflect the movement patterns of residents rather than transient visitors. This distinction is particularly relevant in the context of political districts, which are designed to represent constituents residing within the district boundaries. Following Kang et al. (2020), we infer the population-level human mobility flows between a given origin and destination ($o-d$) CBG pair using the ratio of origin CBG population to the number of SafeGraph mobile devices in the origin. The equation is as follows:

$$\text{Population Flows}_{od} = \text{Device Flows}_{od} \times \frac{\text{Population}_o}{\text{Number of Devices}_o} \quad (1)$$

Using Equation 1, we take the Neighborhood Patterns data set for every month in a year to produce an $o-d$ flows matrix of the average monthly population flows between and within Wisconsin CBGs, where the ith row and jth column intersection is the average monthly population-level flow counts going from the origin CBG to the destination CBG (Figure 2).

To measure social interactions at the CBG level, we use the Facebook Social Connectedness Index (SCI)[1] at the same year. The SCI is calculated based on the number of friendship connections between users in different regions, adjusted for the total number of Facebook users in each area. For two regions i and j, the SCI is defined as:

$$\text{SCI}(Z_i, Z_j) = \frac{\text{Number of Connections}(Z_i, Z_j)}{\text{Users}(Z_i) \times \text{Users}(Z_j)} \quad (2)$$

where Z_i and Z_j represent two different ZIP code tabulation areas (ZCTAs). The SCI reflects the relative probability of a friendship connection between users in Z_i and Z_j compared to other regions, scaled such that the maximum SCI value in the data set is 1,000,000,000 and the minimum is 1. The SCI data are retrieved at the ZCTA level and then converted to an estimate of raw social connections between CBGs. To produce a raw connection value similar to population flows between CBGs, we estimate the

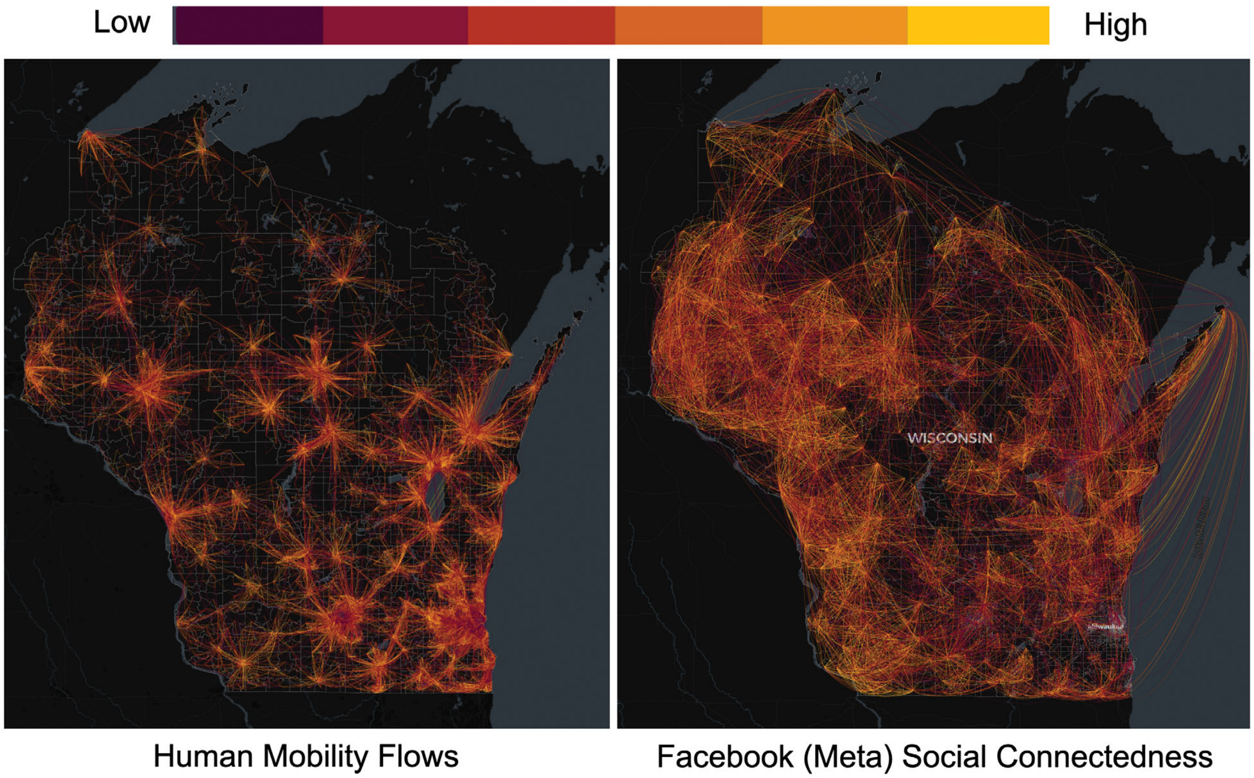

Figure 2. The spatial (human mobility flows) and social interaction (Facebook social connectedness index) data in Wisconsin.

raw number of friendship connections between CBGs, $R(C_{ik}, C_{jl})$, by assuming that the population in a given census unit is proportionate to the number of Facebook users. To relate the fraction of a ZCTA's population residing in a given CBG, we use the Geocorr tool,[2] enabling comparison between spatial and social connections at the CBG level (Figure 2).

Because the provided SCI metric is scaled using an unknown factor, however, the resulting estimates of $R(C_{ik}, C_{jl})$ also carry that same scaling factor. As a result, these estimates allow us to understand the relative distribution of friendship connections between CBGs, but the absolute number of connections cannot be directly interpreted.

Fuzzy Membership

To provide an assessment of each CBG district membership from multiple perspectives, fuzzy membership scores are first calculated based on spatial and social interactions, separately. The membership function $\mu_{D_j}(x)$ quantifies the partial membership of CBG x in district D_j, based on the relative interaction strengths between x and the CBGs in each district. Specifically, the formula is defined as:

$$\mu_{D_j}(x) = \frac{\sum_{y \in D_j} \text{Interactions}_{xy}}{\sum_{i=1}^{n} \sum_{y \in D_i} \text{Interactions}_{xy}}, \tag{3}$$

where $\mu_{D_j}(x)$ represents the membership of CBG x in district D_j, indicating the proportion of interaction strength that x has with all CBGs in district D_j relative to its interactions with CBGs in all districts. Here, D_j denotes a specific district among a set of districts $\{D_1, D_2, ..., D_n\}$, where j represents the index of the district. Using the spatial and social interaction data sets, each CBG is assigned a partial membership in every congressional district based on the strength of its interactions with the CBGs in each district, resulting in one set of fuzzy memberships derived from spatial interactions and another set from social interactions.

To combine the two types of spatial interactions into a combined, interpretable fuzzy membership score, we sought a methodology that would yield results readily explainable to both public and legal audiences. First, in the absence of state-specific guidance regarding the relative importance of each spatial interaction type, we opted for equal weighting. Beyond the weighting of each spatial interaction type, two obvious methods for combining the spatial interactions are addition and multiplication. As agreement between the two data sources indicates a more robust set of community connections, we use multiplication to amplify district membership when strong agreement exists between partial memberships from both perspectives. The combined fuzzy membership for CBG, denoted as $\mu_{D_j}^{\mathrm{combined}}(x)$, is calculated as the element-wise product of the two fuzzy membership sets, followed by renormalization to ensure the partial memberships sum to one. The equation is as follows:

$$\mu_{D_j}^{\mathrm{combined}}(x) = \frac{\mu_{D_j}^{\mathrm{spatial}}(x) \cdot \mu_{D_j}^{\mathrm{social}}(x)}{\sum_{k=1}^{n} \mu_{D_k}^{\mathrm{spatial}}(x) \cdot \mu_{D_k}^{\mathrm{social}}(x)}. \tag{4}$$

As the spatial and social interaction data sets do not cover all of Wisconsin, only CBGs with data from both data sets are included in the analysis. For defuzzification, each CBG is assigned to the district in which it has the highest fuzzy membership. If this assignment differs from the district assignment under the PMC redistricting plan, the CBG is considered *misassigned*. Additionally, for CBGs whose maximum membership district aligns with their PMC assignment but where the maximum fuzzy membership value is less than 50 percent of the CBG's total interactions, the CBG is classified as a *low-membership* CBG. This designation indicates that, although the CBG is placed in the correct district, it has a relatively weak affiliation with that district. This approach can help evaluate a geographic area's district affiliation and their spatial patterns.

Entropy, Information Gain, and Kullback–Leibler Divergence

To assess how dispersed a given CBG's fuzzy membership is across all districts, Shannon information entropy is applied to the fuzzy membership values for each CBG. Specifically, for each CBG x, the entropy is calculated as:

$$H(x) = -\sum_{j=1}^{n} \mu_{D_j}(x) \log_2(\mu_{D_j}(x)), \tag{5}$$

where $\mu_{D_j}(x)$ represents the partial membership of the CBG in district D_j, and n is the total number of districts. To understand the degree to which combination of the two data sets reduces entropy, this calculation is performed separately for the fuzzy membership sets derived from spatial, social, and combined interaction fuzzy memberships.

To evaluate the information gained by combining the perspectives, the information gain is calculated twice: once comparing the spatial perspective to the combined perspective, and once comparing the social perspective to the combined perspective. The information gain when comparing the spatial to the combined perspective is calculated as follows:

$$\text{Information Gain}_{\mathrm{spatial}}(x) = H^{\mathrm{spatial}}(x) - H^{\mathrm{combined}}(x), \tag{6}$$

and the information gain when comparing social to combined is calculated as:

$$\text{Information Gain}_{\mathrm{social}}(x) = H^{\mathrm{social}}(x) - H^{\mathrm{combined}}(x), \tag{7}$$

where $H^{\mathrm{spatial}}(x)$, $H^{\mathrm{social}}(x)$, and $H^{\mathrm{combined}}(x)$ represent the entropy of the CBG's memberships derived from spatial interactions, social interactions, and the combined interactions, respectively. A positive information gain in either case indicates that the combined perspective provides a clearer, less uncertain view of the CBG's district membership compared to the individual spatial or social perspectives.

Kullback–Leibler (KL) divergence is used to quantify the degree to which the social and spatial membership entropy distributions differ from the combined entropy distribution for each district, offering insight into the alignment of each perspective with the integrated view. For a given district D, the KL divergence of the social perspective from the combined perspective is given by:

$$\mathrm{KL}(\mathrm{Social}||\mathrm{Combined}) = \sum_{x \in D} \phi_{\mathrm{Social}}(x) \log \frac{\phi_{\mathrm{Social}}(x)}{\phi_{\mathrm{Combined}}(x)}, \tag{8}$$

where $\phi_{\mathrm{Social}}(x)$ and $\phi_{\mathrm{Combined}}(x)$ represent the normalized distributions of fuzzy membership entropies for each CBG x in the social and combined perspectives, respectively. Similarly, the KL divergence of the spatial perspective from the combined perspective is calculated as:

$$\mathrm{KL}(\mathrm{Spatial}||\mathrm{Combined}) = \sum_{x \in D} \phi_{\mathrm{Spatial}}(x) \log \frac{\phi_{\mathrm{Spatial}}(x)}{\phi_{\mathrm{Combined}}(x)}. \tag{9}$$

To assess the statistical significance of the observed KL divergence, a permutation test is conducted. For each district and entropy type (social or spatial), the observed KL divergence is compared to a distribution generated from 1,000 permuted samples. In each permutation, entropy values within each district are shuffled to remove any geographic relationships, and the KL divergence between this permuted distribution and the combined entropy distribution is calculated. The resulting *p* value represents the proportion of permuted KL divergences that were as extreme or more extreme than the observed value (Good 2000). This test evaluates whether geographic entropy structures significantly influence the information gain within each district. A significant result would indicate that the observed spatial or social distribution aligns with the combined distribution in a nonrandom, geographically patterned way.

Plan Evaluation

Using the data sets and formulas already described, the following analyses are conducted to demonstrate how spatial and social interactions can inform the district memberships of CBGs. In the first analysis, the partial membership of each CBG in its assigned congressional district is examined, allowing for the identification of CBGs that are not strongly affiliated with their assigned district. It further explores the alignment and divergence of COI affiliations based on spatial and social interactions. In the second analysis, we explore the impact of combining spatial and social memberships on incorrectly assigned and low-membership CBGs. By comparing the combined memberships with spatial and social memberships individually, this analysis reveals if spatial or social interactions play a dominant role in defining connections between districts.

Results

Overview of Social, Spatial, and Combined Fuzzy Memberships Across Districts

The maximum fuzzy memberships from spatial, social, and combined perspectives, respectively, are shown Figure 3, where the CBG color corresponds to maximum fuzzy membership value for that CBG. Misassigned CBGs are shown in purple, and low-membership CBGs are shown in tones of white and brown. Overall, mobility-based spatial interactions produced higher maximum membership values, resulting in fewer misassigned CBGs and fewer low-membership CBGs relative to the fuzzy memberships derived from social interactions. These results can be seen in Figure 3, where the left map shows the CBG-level maximum membership values based on social interactions (Facebook friendships), and the middle map shows maximum membership values based on spatial interactions (human mobility flows).

The combination of social and spatial fuzzy memberships through element-wise multiplication and normalization generates a fuzzy membership set that

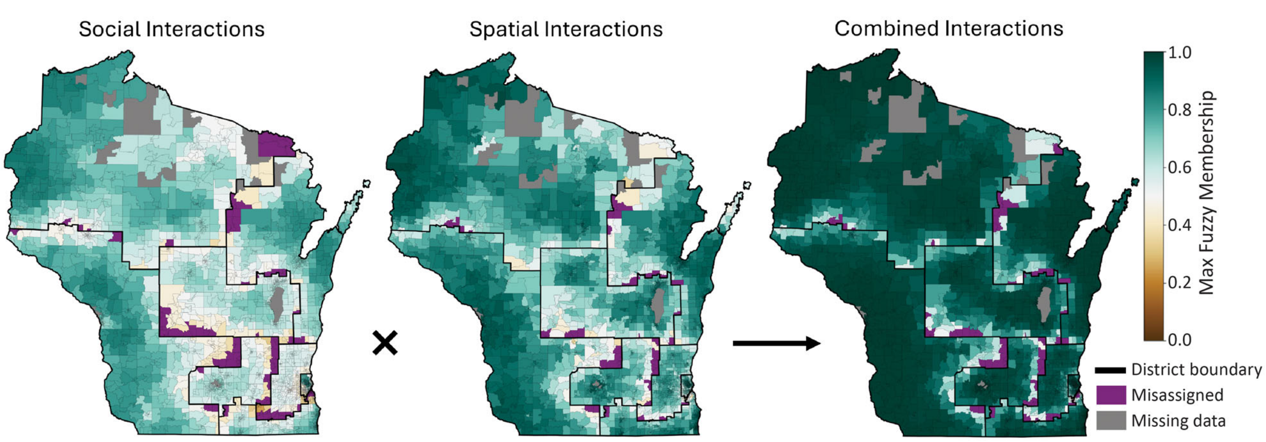

Figure 3. Maximum fuzzy membership value for each census block group (CBG), from social, spatial, and combined interactions. Element-wise multiplication of fuzzy memberships for each CBG produces much higher maximum membership values for most CBGs, showing that the spatial and social interaction data sets tended to agree on their maximum membership districts.

increases the membership value where both spatial and social perspectives align, while diminishing the value where they do not. As seen in the right-most map in Figure 3, combination results in more CBGs achieving a maximum fuzzy membership value of close to 1 compared to the social and spatial interaction maximum partial membership maps on the left. Additionally, there are significantly fewer misassigned and low-membership CBGs in this combined approach. Table 1 illustrates this, with the misassigned and low-membership counts shown in bold if they are less than or equal to the corresponding counts from the social perspective, and underlined if they are less than or equal to the corresponding counts from the spatial column. Generally, there is an averaging effect across districts for the number of misassigned CBGs. The combined number of misassigned CBGs typically falls between the counts observed from the social and spatial perspectives. It is usually closer to the count from the spatial perspective, however, suggesting that fuzzy memberships based on spatial interactions tend to show stronger affiliations with their primary district. In contrast, social interaction-based fuzzy memberships appear more dispersed, indicating weaker dominant affiliations.

For low-membership counts, all districts except one show a substantial decrease after combining both social and spatial perspectives. This trend supports the idea that, in most cases, the majority membership district for each CBG aligns across their residents' social and spatial dimensions. Regarding the locations of misassigned and low-membership CBGs from all perspectives, CBGs located along district borders are the most likely to be misassigned or low membership. Indeed, this pattern is evident in all three maps in Figure 3, where purple CBGs indicate misassigned CBGs, and white and brown CBGs represent low-membership CBGs. Similarly, CBGs that are farthest away from other districts tend to have the highest membership, although this is not always the case.

When taken as a whole, the more focused memberships from the spatial perspective, along with the tendency of misassigned and low-membership CBGs to be along district borders, likely reflects the role of distance in the underlying spatial interaction networks. Although people are still more likely to form friendships with people who are geographically close (Bailey et al. 2020), social interactions are generally less constrained by distance compared to mobility-based physical interactions. Additionally, the social interaction network used here reflects the number of Facebook friendships between CBGs. As friendships are generally established and then maintained, they do not require the cost of distance to be paid regularly. In contrast, spatial interactions (e.g., a trip between CBGs) require an investment of time, money, and more, every time the trip is made. As such, people in one CBG can maintain many friendships with those in distant CBGs at a relatively low cost, whereas the investment of physically moving between such CBGs is much greater. Accordingly, it would be expected that the fuzzy memberships of CBGs based on social interactions would be more dispersed geographically, whereas those based on spatial interactions would be much more localized. This is indeed reflected in the misassigned and low-membership counts of each respective category.

Table 1. Counts and percentages of misassigned and low-membership census block groups by district

	Social		Spatial		Combined	
District	Misassigned	Low-membership	Misassigned	Low-membership	Misassigned	Low-membership
1	98 (16.4%)	59 (9.9%)	11 (1.8%)	27 (4.5%)	**58** (9.7%)	***12*** (2.0%)
2	15 (3.1%)	37 (7.6%)	11 (2.3%)	6 (1.2%)	***2*** (2.5%)	***1*** (0.2%)
3	24 (4.2%)	65 (11.3%)	9 (1.6%)	17 (3.0%)	**12** (2.1%)	***0*** (0.0%)
4	2 (0.3%)	236 (33.6%)	0 (0.0%)	10 (1.4%)	***0*** (0.0%)	**1** (0.1%)
5	24 (4.5%)	192 (36.0%)	10 (1.9%)	21 (3.9%)	**15** (2.8%)	***5*** (0.9%)
6	20 (3.3%)	72 (11.8%)	10 (1.6%)	13 (2.1%)	**13** (2.1%)	***0*** (0.0%)
7	12 (2.0%)	17 (2.8%)	5 (0.8%)	7 (1.1%)	**7** (1.1%)	***0*** (0.0%)
8	17 (3.2%)	28 (5.2%)	11 (2.0%)	9 (1.7%)	**15** (2.8%)	***0*** (0.0%)

Note: Values in the combined column are bold if they are less than or equal to the corresponding social values and italic if they are also less than or equal to the corresponding spatial values, indicating stronger alignment or lower counts in the combined metrics compared to the individual social and spatial measures.

Fuzzy Memberships in Districts 1 and 4

To further demonstrate how the spatial, social, and combined fuzzy memberships can be used to understand district affiliation in local communities, a deep examination of Districts 1 and 4 is provided (see Figure 1 for district reference on the map). From the social interaction network perspective, District 1 had the highest number of misassigned CBGs, at ninety-eight (16.4 percent of the CBGs in that district), compared with two (0.3 percent) misassigned CBGs in District 4. The majority of misassigned CBGs in District 1 are located within the Milwaukee city boundary, and of the misassigned CBGs in District 1 that fall within Milwaukee's boundary, 100 percent have their highest fuzzy membership in District 4. This indicates that the social interactions of those CBGs are strongly directed within the city of Milwaukee, rather than without.

Looking at low-membership CBGs reveals a different story, however. From a social network perspective, District 4 has only two misassigned CBGs but includes 236 (33.6 percent) low-membership CBGs, suggesting that many CBGs in this district exhibit highly divided membership between two or more districts. Most of the low-membership CBGs in District 4 are located on the southeastern side of the district, close to Districts 1 and 5, suggesting a large degree of social community overlap between these regions. Indeed, looking at the second maximum fuzzy membership district for low-membership CBGs in District 4 that are outside of the city of Milwaukee, 96 percent of them are District 5. From the social perspective alone, such low-membership CBGs could likely be placed in either District 4 or 5, as their membership is relatively evenly dispersed between CBGs in those two districts.

Focusing now on the combined membership values for District 1 (Table 1), it is clear that combining the social and spatial fuzzy memberships has a balancing effect on the number of misassigned CBGs. Compared to using only social connections, which results in ninety-eight misassigned CBGs, the combined misassigned CBG count decreases to fifty-eight. Conversely, the combined count of fifty-eight is higher than the number of misassigned CBGs when only the spatial perspective is used. In terms of low-membership CBGs, the combined approach yields twelve, which is significantly lower than for either the social or spatial perspectives alone. Looking at the low-membership CBGs (from the spatial perspective) in District 1, all twenty-seven became misassigned after combination with the social perspective, showing how combining the two perspectives is most likely to change the district assignment of CBGs that have relatively weak membership in their majority district from a given perspective. In terms of location, the twenty-seven CBGs in District 1 that go from low membership to misassigned after combination are primarily located near the southern border of Milwaukee (Figure 4). Low-membership CBGs (indicated by white and tan colors) in the spatial fuzzy membership map align with low-membership and misassigned CBGs in the social perspective. The geographic clustering, along with fact that all twenty-seven CBGs of the misassigned CBGs in District 1 have their majority membership in District 4, strongly suggests that this area corresponds to a community that is deeply tied to the rest of the city of Milwaukee and District 4. In terms of the practical importance of these findings to redistricting, these patterns suggest that these CBGs from District 1 might be among the first candidates for district reassignment if the PMC map were to be redrawn.

Despite instances of misalignment, many CBGs within Districts 1 and 4 reach nearly full membership (a fuzzy membership value of 1.0) in their designated districts after combination. Indeed, District 4 initially had the highest number of low-membership CBGs (236) from the social perspective alone. After combining perspectives, however, both the misassigned and low-membership counts for District 4 dropped to zero, underscoring the extent of agreement between the data sets and illustrating how uncertainty in one perspective can be mitigated through alignment with the other.

Election Outcomes for Districts 1 and 4

As shown in the analysis of CBGs in District 1 that fall within the border of the city of Milwaukee, there is strong geographic clustering in the district affiliation of CBGs near district and city lines. Because such geographically based communities might also be reflected in political affiliation, the 2020 presidential election voting outcomes for precincts in Districts 1 and 4 are briefly examined. Figure 4 presents the fraction of votes cast for Biden (the Democratic candidate) as a percentage of the total votes cast for Biden and Trump

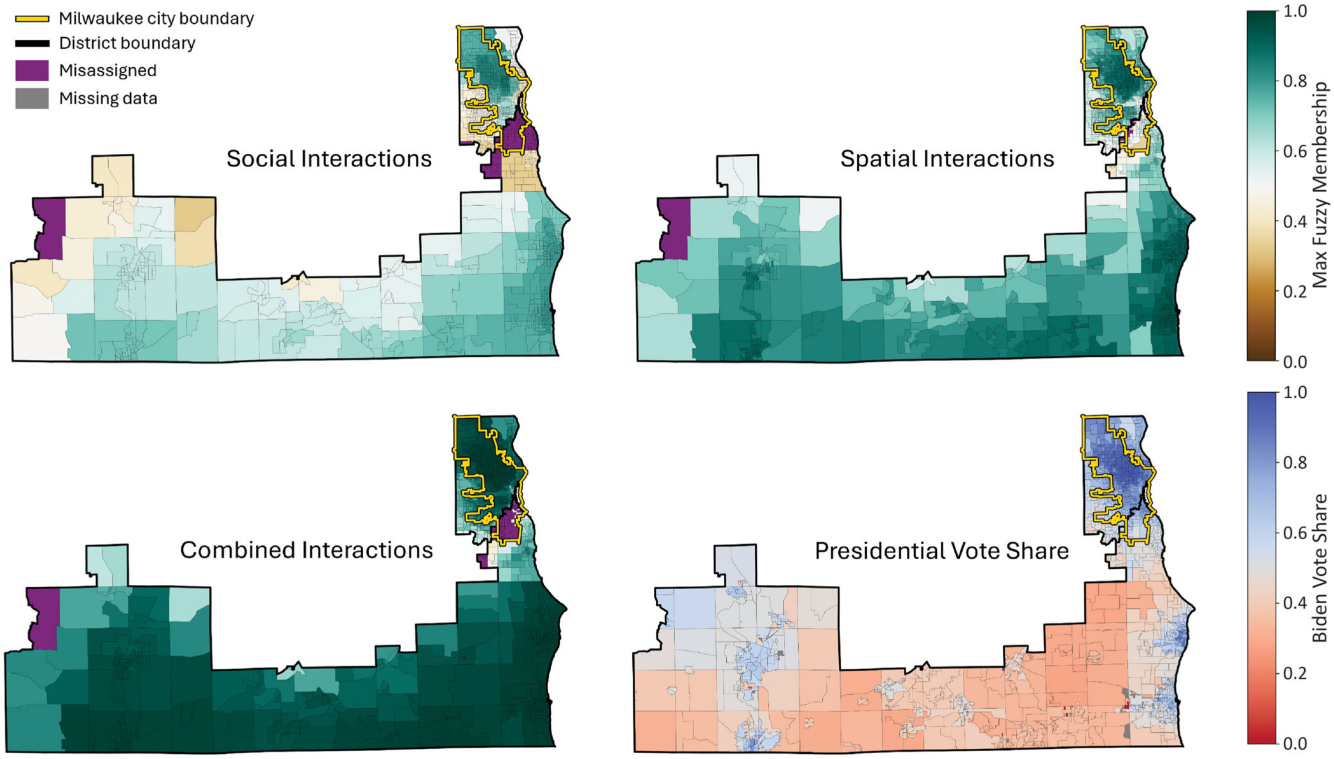

Figure 4. Cutaway maps focusing on Districts 1 (lower district) and 4 (higher district) illustrate different perspectives on census block group–level interactions and affiliations: (A) social interactions (top left), (B) spatial interactions (top right), (C) combined interactions (bottom left), and (D) Biden's vote share by precinct in the 2020 election (bottom right).

(the Republican candidate) within these precincts. This map reveals that the majority of precincts in central Milwaukee exhibit high voting percentages for Biden, with precincts in the city center having an average 86 percent vote share for him.

Overall, the combined fuzzy memberships offer a multifaceted perspective on the COI connections that district boundaries aim to encapsulate, enabling the identification of areas where district boundaries place CBGs in districts with which they are not most strongly affiliated. Although the number of misassigned CBGs decreases or remains the same using the combined approach, the CBGs that are still misassigned after combination would be the first to evaluate for district reassignment when developing a new redistricting plan, as such CBGs have the strongest ties to communities outside of their currently assigned district. Likewise, CBGs with low membership in their current district could be evaluated for reassignment, given that they only share weak connections within their currently assigned district.

In terms of how this type of analysis could be formalized as a framework for analyzing redistricting plans, one approach might be to evaluate the number of misassigned CBGs in each district and compare these numbers across districts. If certain districts have many more misassigned CBGs than others, these outliers could be addressed first in developing a new plan. For instance, using the combined partial memberships, District 1 has fifty-eight (9.7 percent) misassigned CBGs and twelve (2.0 percent) low-membership CBGs, which is much higher than the best-performing district, District 4, which has no misassigned CBGs and one low-membership CBG. The large differences in the number of misassigned and low-membership values between the worst district and the rest provides a clear starting point for evaluating which district boundaries should be reconsidered. In the context of legal challenges to a congressional district plan, plan statistics for misassigned and low-membership CBGs could be compared to the distribution of those same statistics for an ensemble of plans, which we explore later.

Entropy Analysis of CBGs

The previous sections explore how the combination of two data sources can be used to evaluate CBG district affiliation, focusing primarily on how this combination affects the value of the maximum fuzzy membership. Although maximum fuzzy membership is useful for assessing CBG district affiliation (e.g., identifying misassigned and low-membership CBGs), this approach can obscure the extent to which correctly assigned CBGs have divided memberships across multiple districts. To address this, entropy—a measure of class dispersion—is used here to analyze the extent to which CBGs exhibit divided memberships. In the context of evaluating CBG district membership from multiple perspectives, information gain can reveal how combining spatial and social perspectives can reduce membership ambiguity across all districts in which a CBG has partial membership. By comparing entropy from spatial and social perspectives, independently, to the entropy of the combined perspective, the information gain for each CBG can be calculated, offering insight into how much clearer or less dispersed each CBG's membership becomes after integrating the data sources.

As shown in Figure 5, the combined fuzzy membership map (upper right) exhibits significantly lower entropy than either the individual social or spatial maps. The maps in the lower part of Figure 5 illustrate the changes in entropy when transitioning from the social and spatial perspectives to the combined perspective, both of which demonstrate a significant decrease in entropy, reflecting information gain. The results indicate that all CBGs, from both spatial and social perspectives, experienced

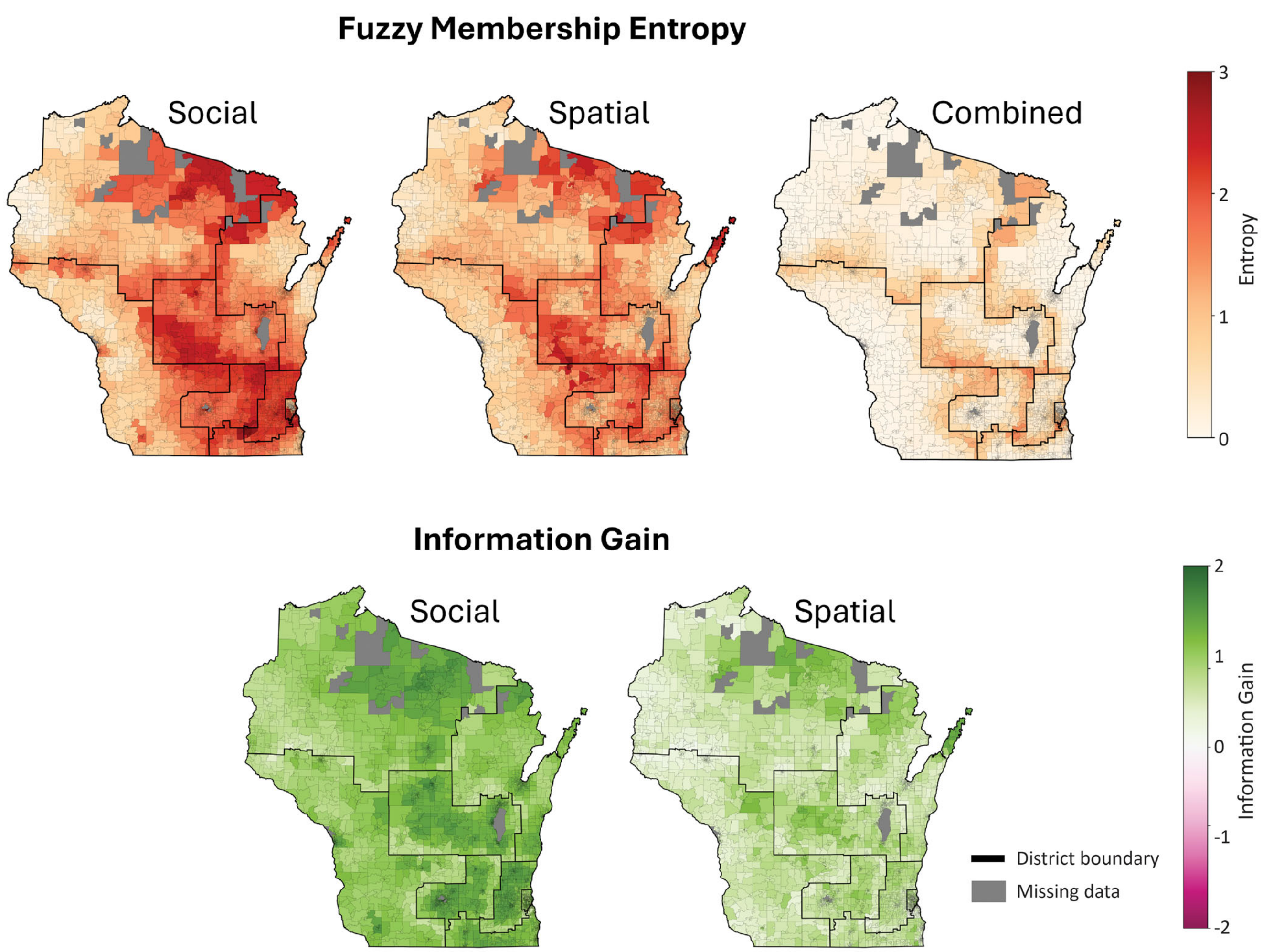

Figure 5. Fuzzy membership entropy and information gain for social, spatial, and combined perspectives in Wisconsin congressional districts in the People's Map Commission proposal. The top row displays the entropy of fuzzy membership values from social (left), spatial (center), and combined (right) data sets.

information gain from combination. Overall, the social interaction perspective saw the greatest decrease in entropy from combination, with the spatial interaction perspective showing much more modest decreases in entropy. Generally, CBGs nearest to district borders have the highest entropy, with combination producing the most information gain in the district interiors.

District-Level Entropy Distributions

In the context of evaluating a given district proposal, it might be desirable to compare how different districts perform in terms of divided membership. This would provide a quantitative measure of the extent to which all geographic subunits (e.g., CBGs) within each district exhibit divided membership. Such a measure could serve as a starting point for identifying districts that are most in need of modification to better align their boundaries with underlying COIs. Accordingly, the entropy distribution of the combined fuzzy memberships in each district is shown in Figure 6. Overall, District 5, which covers suburban and rural areas to the west and north of the city of Milwaukee, has relatively high entropy among its CBGs, in terms of the district median and interquartile range. The high entropy values in District 5 reflect more divided membership across the district, with the majority of high-entropy CBGs located along district borders with neighboring districts. This aligns with the misassigned and low-memberships counts for District 5 in Table 1, as District 5 has the second highest count in each of the categories. Conversely, District 8 shows a distribution of entropy values skewed lower than those of other districts, suggesting that many of its CBGs are well-placed within the district. This might also be influenced by the fact that much of District 8 borders Lake Michigan and the state of Minnesota, rather than another congressional district, as CBGs along district borders tend to have higher entropy values. This factor alone, though, does not fully explain District 8's lower entropy, as District 7, which borders Lake Superior and has the Mississippi River as its entire western boundary, shows an entropy distribution more in line with other districts. District 1 exhibits the widest interquartile range, indicating significant variability in divided membership within the district. This suggests that some regions within District 1 have considerably higher levels of divided membership, whereas others are more cohesively situated within the district. The large spread suggests potential outlier areas with pronounced divided memberships, compared to regions that are well-aligned within the district. Many of the highest entropy CBGs in District 1 are located

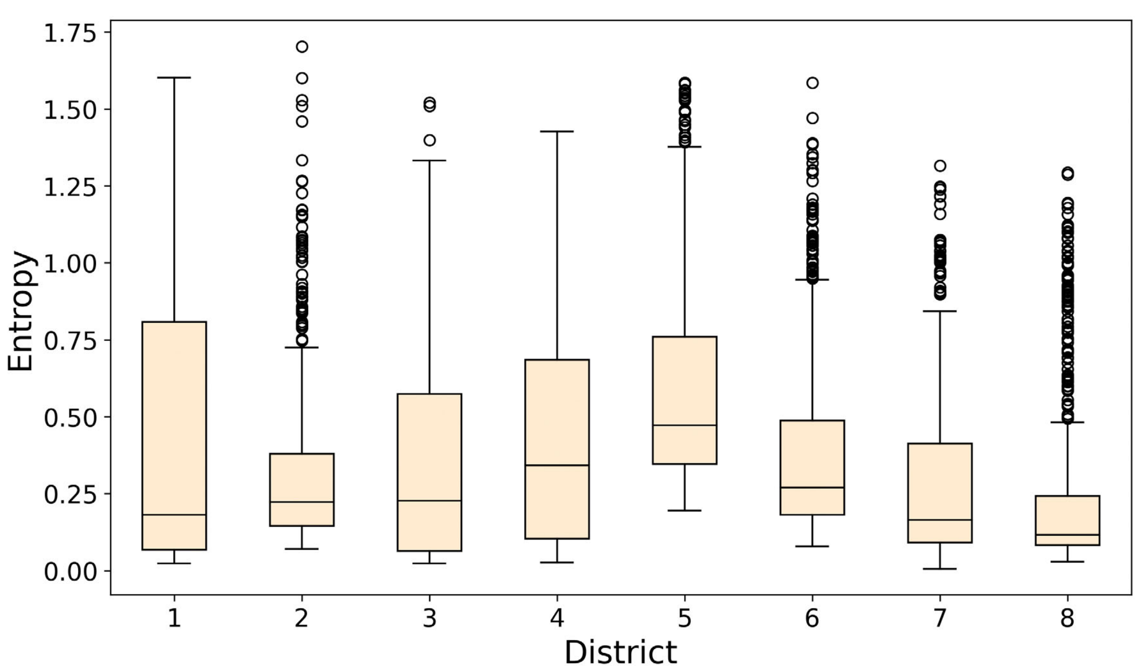

Figure 6. Box plots of entropy values for each congressional district, showing the distribution of dispersed membership values among census block groups within each district.

near the border of District 4, reflecting the high count of low-membership and misassigned CBGs in District 1 near its border with District 4, as discussed in the previous section (see Figure 4).

KL Divergence

To quantitatively evaluate the changes in entropy at the district level, the KL divergence for each district is plotted in Figure 7. A higher KL divergence indicates a greater difference between the individual (social or spatial) entropy distributions and the combined distribution, suggesting that one of the perspectives (social or spatial) is more distinct from the combined membership structure within that district. In all districts, the KL divergence shows that the difference between the social membership entropy distribution and the combined membership entropy distribution (blue bars) is greater than the difference between the spatial membership entropy distribution and the combined distribution (red bars). This suggests that the spatial membership entropy distribution is more closely aligned with the combined membership entropy distribution compared to the social membership entropy distribution. In Districts 4 and 5, the differences between social and spatial divergences are small, indicating a relatively balanced influence from both perspectives on the combined fuzzy memberships.

Overall, the social membership entropy tends to diverge more from the combined membership entropy. This, along with the lower entropy across virtually all of the combined membership CBGs, shows that the social fuzzy memberships generally saw greater information gain than the spatial fuzzy memberships, a sign that social fuzzy memberships tend to be more dispersed, and spatial fuzzy memberships less dispersed. As the goal of combining the spatial and social perspectives is to build a multiperspective evaluation of CBG district affiliation, a key component to consider is if the combination of perspectives increases confidence in the calculated affiliation or not. As all CBGs, from both perspectives, saw decreases in entropy after combination, we can have more confidence that the combined fuzzy membership values reflect an aligned perspective on COIs. The KL divergence test then quantifies the extent of the information gain and helps us understand which data source contributes most to the increase in fuzzy membership alignment.

The global reduction in entropy from combination does not necessarily imply alignment with the correct district; rather, it reflects a less divided

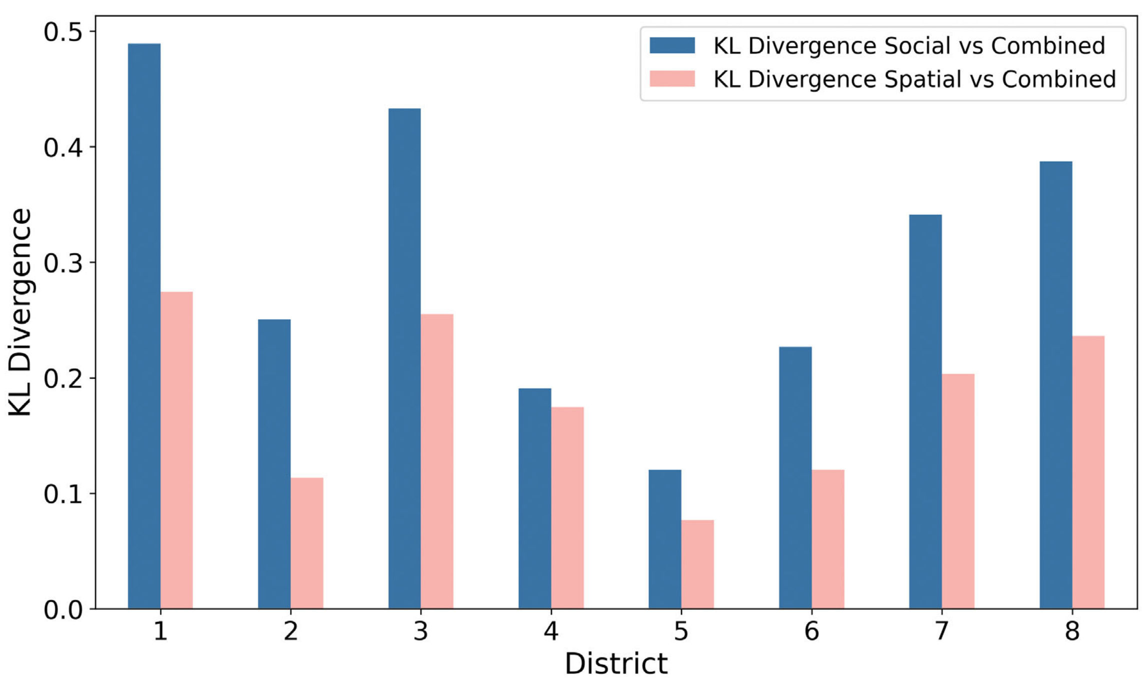

Figure 7. Comparison of Kullback–Leibler (KL) divergence between social, spatial, and combined membership entropy distributions across districts.

membership distribution across multiple districts, pointing to strong agreement between the two types of interactions, albeit with some regional variation. In particular, the areas with smallest information gain also tend to be the areas with the highest entropy to begin with—CBGs near district borders. The overall agreement between the spatial and social perspectives, however, gives us confidence that the spatial and social interactions are capturing different perspectives of the same underlying interaction communities.

To evaluate the statistical significance of the KL divergence results, a permutation test is conducted. The null distribution represents what the entropy differences would look like if there were no real geographic structure or alignment between spatial, social, and combined distributions; that is, if the combination of the two perspectives was not causing reductions in entropy in particular CBGs. Table 2 summarizes the p values for each district, indicating whether the divergence between the individual social or spatial membership entropy distributions and the combined membership entropy distribution is statistically significant.

The results highlight district-specific patterns in how social and spatial perspectives align with the combined entropy structure. Districts 1, 4, and 8 exhibit a significant divergence between the social membership entropy distribution and the combined membership entropy distribution ($p < 0.05$), suggesting that social memberships experienced substantial changes in entropy when combined with spatial data. Notably, District 8 shows significant divergence for both the social and spatial perspectives, indicating substantial changes in both distributions upon combination. This suggests that, for District 8, the social and spatial perspectives differ markedly from each other, leading to a distinct combined structure. Conversely, Districts 2, 5, 6, and 7 show no significant divergence for either the social or spatial perspectives, implying that both distributions closely align with the combined distribution. This lack of significant divergence suggests that, for these districts, the combined map integrates both perspectives without substantial alteration, maintaining the consistency of the social and spatial input maps. Notably, all four of these districts exhibit information gain, as seen in Figure 5, but the information gain shows low variability across CBGs within each district. This uniformity suggests that entropy changes, although present, are consistent across the district and do not form concentrated spatial clusters, leading to an overall alignment of spatial or social perspectives across the combined map. Together, these findings indicate that the combined membership entropy distribution often aligns more closely with the spatial perspective. This closer alignment suggests a stronger influence from spatial fuzzy membership values, indicating that those values were generally less dispersed.

Table 2. p values for Kullback–Leibler divergence of social and spatial membership entropy distributions relative to the combined membership entropy distribution in each district

District	Social	Spatial	Significant divergence from combined
1	0.00	1.00	Social
2	1.00	1.00	None
3	0.92	1.00	None
4	0.00	1.00	Social
5	1.00	1.00	None
6	1.00	1.00	None
7	1.00	1.00	None
8	0.00	0.00	Both

Fuzzy Membership Ensemble Evaluation

The evaluation of district plans using metrics such as fuzzy membership entropy naturally leads to the topic of plan optimization, as these metrics can serve as objective functions in optimization algorithms. The inherent computational complexity of the redistricting problem (NP-hard), however, renders the determination of a globally optimal solution computationally intractable for most real-world political redistricting scenarios (Altman 1997). Therefore, as a methodological shift, outlier analysis has emerged as a prevalent analytical framework for evaluating district plans. This approach assesses the quality of a proposed plan by comparing it against an ensemble of diverse, valid alternative configurations (Duchin 2018; Herschlag, Dibaeinia, and Bonica 2022).

This method has been effectively used in legal challenges to state-level redistricting plans and is expected to play an increasingly prominent role in redistricting litigation (Ramachandran and Gold 2018). In this study, we apply outlier analysis to evaluate the quality of the PMC map by comparing its fuzzy membership metrics against a sample distribution of 10,000 district plans generated using the ReCom algorithm (DeFord, Duchin, and Solomon 2019). Each plan in this ensemble is analyzed using combined fuzzy memberships, with three key metrics calculated across all districts: (1) the sum of entropy, (2) the count of misassigned CBGs, and (3) the count of low-membership CBGs. The PMC plan is then compared to this distribution to identify whether it exhibits outlier characteristics, providing valuable insights into its reasonableness with respect to these measures (Figure 8).

Using a two-tailed *t* test, the results indicate that the PMC map is a significant outlier for both the sum of entropy and misassigned CBG metrics, with two-tailed *p* values of 0.0009 and 0.0004, respectively. This suggests that these values are unusually low compared to the distribution of values across all assignments. The low-membership CBGs metric, however, does not show the PMC map to be a significant outlier, with a *p* value of 0.3634, indicating that this value is more typical within the distribution. For the metrics evaluated here, being an outlier on the low end of the distribution indicates that the PMC map performs better than the vast majority of plans, meaning it produces fewer misassigned CBGs and a lower sum of entropy compared to most plans. Conversely, an outlier with a high value would indicate that the map performs worse than the majority of plans. Overall, the outlier analysis helps gives context to the fuzzy membership analysis of the PMC plan. One aspect of the fuzzy membership results that particularly benefits is the location of the misassigned, low-membership, and high entropy CBGs, which tend to be located along district boundaries. Looking at the PMC map by itself, we might believe that district boundaries are unnecessarily splitting interaction communities, given the relatively high incidence of misassigned, low-membership, and high-entropy CBGs along the boundaries. Any imposition of an artificial political boundary on top of the underlying interaction networks, however, would produce district assignments where boundary CBGs are the most likely to be have highly divided memberships, given the distance decay effect in spatial and social interactions. Therefore, comparison with a distribution of alternative maps helps us understand if the boundary effects present in the PMC map are within a normal range, or if they are indeed outliers in one way or another.

Discussion

As noted previously, CBGs without spatial or social interaction data were excluded from the analysis. Although this limitation is less critical in a research context, its impact could be significantly more pronounced if the proposed methodologies were applied in a legal setting to evaluate voting districts. To address the issue of missing data, various approaches can be considered. These include imputation techniques to estimate missing values, sensitivity analyses to assess the impact of missing data, or incorporating supplementary data sources to improve coverage. A common method is to estimate missing values, often using spatially weighted nearest-neighbor estimates in geographical analyses. As seen in the social and spatial interaction patterns for District 7 in Figure 3, however, neighboring CBGs can sometimes have very different memberships. This might occur if one CBG represents a

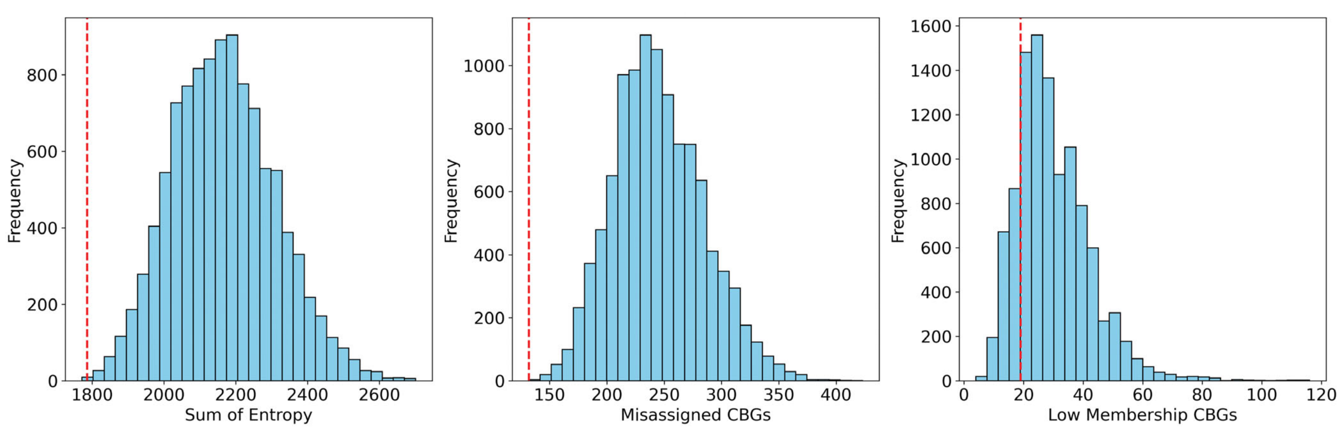

Figure 8. Outlier analysis of People's Map Commission (PMC). Compared to the ReCom distribution, the PMC values for the sum of entropy and misassigned census block groups (CBGs) count were outliers, whereas the low-membership CBG count was not.

town with strong connections to other districts (e.g., as a transportation hub), whereas adjacent CBGs consist of local residents with primarily local interactions. In the context of redistricting, CBGs that are centrally located within a district are less likely to change districts, so missing data are less critical for these CBGs. For CBGs near district borders, however, estimates of interactions might need to be made more carefully. If data are missing from only one source—such as if a CBG has spatial interaction data but lacks social interaction data—using the available data might suffice, especially given the high degree of overlap between the two estimates. Alternatively, other data sources can be used to fill gaps when more critical data sets, like human mobility flows, which very closely reflect general human activity, are unavailable.

Regarding data representativeness, both interaction data sets in this research were collected in one year. To derive more stable estimates of district affiliations based on spatial and social interactions, we recommend using long-term data sets spanning several years. Unfortunately, this was not feasible in our case due to social media data availability constraints.

Given that this work evaluates political districts, it is natural to wonder why political data, such as voting patterns, were not employed in the fuzzy membership calculations. To explain this, we note that term *political redistricting* refers to the political nature of the process and outcomes—rather than implying that districts should be drawn based on partisan affiliation. As described earlier, district boundaries are typically grounded in nonpartisan principles such as preserving COIs and ensuring compactness. Furthermore, COIs are generally not defined in terms of political affiliation, either. In terms of redistricting law, states vary widely in how they permit the use of political data (e.g., voting history or party registration): Some prohibit its use entirely (National Conference of State Legislatures 2021a), others restrict its use to neutral evaluation (National Conference of State Legislatures 2021c), and a growing number explicitly require it to assess partisan fairness or competitiveness (National Conference of State Legislatures 2021b). Rather than use political data to measure region affiliation, we instead focus on measuring region affiliation in terms of community interactions, which we believe is more in line with the literal definition of COIs as communities with interests in common.

Finally, we briefly highlight how this type of analysis could be applied to understanding contested regions. Contested regions emerge when two or more groups claim the same piece of land as belonging to them, often leading to conflict. Although a variety of factors, such as political, ethnic, geographic, religious, and historical considerations, contribute to these disputes (Murphy 2004; Schultz 2017), the use of fuzzy membership methodology with spatial interaction networks to analyze regional affiliations could offer a present-day perspective on the divided affiliations of such places. By providing a detailed description of the contemporary affiliations and interactions within these regions, this approach could offer valuable insights into the modern communities that engage with and use these contested spaces.

Conclusion

In this work, we present a spatial interaction-based fuzzy membership framework to provide subdistrict level evaluations of region affiliation, where multitype interactions represent the various COI connections that a given geographic area has with other areas in each district. To more comprehensively model the spatial interaction communities that constitute COIs in this work, we draw on both the social sensing literature and the splatial framework described by Shaw and Sui (2020) to justify the use of multiple interaction types, ultimately employing both spatial and social interaction types.

The combination of spatial-social interactions helps focus fuzzy membership when they are aligned, and increase membership dispersion when they are not, allowing for a more nuanced understanding of the interactions that underlie COIs across districts. By evaluating regional affiliation at the CBG level, this work introduces a novel method to assess whether a given CBG is assigned to the most appropriate district based on the strength of its connections with other CBGs—a perspective currently underexplored in redistricting literature. The proposed methodology is first applied to evaluate a given plan independently, and then extended through outlier analysis, demonstrating how it can be integrated into existing methods of redistricting analysis. The use of fuzzy memberships in relation to political boundaries also illustrates how the proposed methodology can be applied to understand border regions more broadly, a topic that is highly relevant in fields such as geography, political science, and other regionalization-related disciplines.

Several future research areas could be explored. Although spatial and social interaction data sets were employed to capture various types of human interaction, additional data sets could be explored for use in measuring COI connections. Additionally, the modeling of COIs with multitype interactions could be compared with the COIs that are reported by constituents. If a constituent claims that some set of areas are part of the same COI, this claim could be quantitatively compared against the interaction strengths for those same areas. This would also provide a more objective way of deciding on which COIs to prioritize when multiple, overlapping COIs are identified and contested by constituents.

Furthermore, the proposed methodology can be extended to evaluate COIs in relation to natural boundaries such as rivers, lakes, or mountain ranges. Intuition might suggest that communities separated by these features are weakly connected, but our quantitative spatial interaction measures offer a means to objectively assess the actual strength of connections between such communities. This approach allows for a direct comparison of intuitive assumptions against empirical interaction data, providing a more robust understanding of COI formation in the presence of natural barriers. In selecting spatial interactions for this study, we used human mobility flows and social network connections to capture a broad range of interactions relevant to communities. The methodology is adaptable to any interaction type deemed relevant to a specific COI definition, however, whether established by states or researchers.

This research contributes to the redistricting literature by developing novel approaches to modeling COIs and assessment of region affiliation using spatial and social interactions. This use of fuzzy membership on spatial interactions also contributes to the geography and GIScience literature on region boundaries, providing a methodological framework for quantifying the divided regional affiliation of vague or contested boundary areas.

Disclosure Statement

No potential conflict of interest was reported by the authors.

ORCID

Jacob Kruse https://orcid.org/0000-0001-7752-3190
Song Gao https://orcid.org/0000-0003-4359-6302

Notes

1. See https://dataforgood.facebook.com/dfg/tools/social-connectedness-index/.
2. See https://mcdc.missouri.edu/applications/geocorr.html.

References

Ahmed, W., Q. Tan, Y. A. Solangi, and S. Ali. 2020. Sustainable and special economic zone selection under fuzzy environment: A case of Pakistan. *Symmetry* 12 (2):242. doi: 10.3390/sym12020242.

Altman, M. 1997. The computational complexity of automated redistricting: Is automation the answer. *Rutgers Computer & Technology Law Journal* 23:81.

Aydin, O., M. V. Janikas, R. M. Assunção, and T.-H. Lee. 2021. A quantitative comparison of regionalization methods. *International Journal of Geographical Information Science* 35 (11):2287–315. doi: 10.1080/13658816.2021.1905819.

Bae, C. J.-H., and D. R. Montello. 2018. Representations of an urban ethnic neighbourhood: Residents' cognitive boundaries of Koreatown, Los Angeles. *Built Environment* 44 (2):218–40. doi: 10.2148/benv.44.2.218.

Bailey, M., P. Farrell, T. Kuchler, and J. Stroebel. 2020. Social connectedness in urban areas. *Journal of Urban Economics* 118:103264. doi: 10.1016/j.jue.2020.103264.

Bennett, B. 2001. What is a forest? On the vagueness of certain geographic concepts. *Topoi* 20 (2):189–201. doi: 10.1023/A:1017965025666.

Bezdek, J. C. 1981. *Pattern recognition with fuzzy objective function algorithms*. New York: Springer Science & Business Media.

Brown, D. G. 1998. Mapping historical forest types in Baraga County Michigan, USA as fuzzy sets. *Plant Ecology* 134 (1):97–111. doi: 10.1023/A:1009796502293.

Burrough, P. A., and A. U. Frank. 1996. Natural objects with indeterminate boundaries. *Geographic Objects with Indeterminate Boundaries* 2:3–28.

Burrough, P. A., and R. A. McDonnell. 1989. Fuzzy classification methods for determining land suitability from soil profiles. *Journal of Soil Science* 40 (3):477–91.

Cannon, S., M. Duchin, D. Randall, and P. Rule. 2022. Spanning tree methods for sampling graph partitions. *arXiv Physics*:2210.01401.

Cao, R., W. Tu, C. Yang, Q. Li, J. Liu, J. Zhu, Q. Zhang, Q. Li, and G. Qiu. 2020. Deep learning-based remote and social sensing data fusion for urban region function recognition. *ISPRS Journal of Photogrammetry and Remote Sensing* 163:82–97. doi: 10.1016/j.isprsjprs.2020.02.014.

Chavoshi, S., W. N. A. Sulaiman, B. Saghafian, M. N. B.

Sulaiman, and L. Abd Manaf. 2013. Regionalization by fuzzy expert system based approach optimized by genetic algorithm. *Journal of Hydrology* 486:271–80. doi: 10.1016/j.jhydrol.2013.01.033.

Chen, S. J., S. S.-H. Wang, B. Grofman, R. F. Ober, Jr., K. T. Barnes, and J. R. Cervas. 2022. Turning communities of interest into a rigorous standard for fair districting. *Stanford Journal of Civil Rights and Civil Liberties* 18 (1):101–89.

Chen, W., H. Huang, J. Dong, Y. Zhang, Y. Tian, and Z. Yang. 2018. Social functional mapping of urban green space using remote sensing and social sensing data. *ISPRS Journal of Photogrammetry and Remote Sensing* 146:436–52. doi: 10.1016/j.isprsjprs.2018.10.010.

Chen, Y., Z. Zhang, and T. Liang. 2019. Assessing urban travel patterns: An analysis of traffic analysis zone-based mobility patterns. *Sustainability* 11 (19):5452. doi: 10.3390/su11195452.

Chen, Z., Y. Zou, and Y. Tan. 2025. How far to go to encounter the differences: Examining the nonlinear relationships between low-income migrants' activity-space segregation and traveling distances in Shenzhen, China. *Annals of the American Association of Geographers* 115 (7):1651–73. doi: 10.1080/24694452.2025.2504584.

Cohn, A. G., and N. M. Gotts. 1996. The "egg-yolk" representation of regions with indeterminate boundaries. In *Geographic objects with indeterminate boundaries*, ed. P. A. Burrough and A. U. Frank, 171–87. London and New York: Taylor & Francis.

Contisciani, M., C. De Bacco, and A. Braunstein. 2020. Extracting overlapping communities from multilayer networks. *Physical Review E* 101 (6):062311.

de Cobos-Silva, S. G., M. Gutiérrez-Andrade, E. Rincón-García, R. Mora-Gutiérrez, P. Lara-Velázquez, and A. Ponsich. 2017. Fuga, a fuzzy greedy algorithm for redistricting in Mexico. *Fuzzy Economic Review* 22 (2):2268. doi: 10.25102/fer.2017.02.01.

DeFord, D., M. Duchin, and J. Solomon. 2019. Recombination: A family of Markov chains for redistricting. *arXiv preprint*:1911.05725.

Ding, C., J. Tang, Z. Tang, M. Deng, W. Wang, and H. Liu. 2025. Geographical scene: The natural unit for geographical analysis and its recognition based on data with spatial and semantic features. *Annals of the American Association of Geographers*. Advance online publication. 10.1080/24694452.2025.2511945.

Dong, H., M. Wu, X. Ding, L. Chu, L. Jia, Y. Qin, and X. Zhou. 2015. Traffic zone division based on big data from mobile phone base stations. *Transportation Research Part C* 58:278–91. doi: 10.1016/j.trc.2015.06.007.

Duchin, M. 2018. Outlier analysis for Pennsylvania congressional redistricting. Technical report, Tufts University, Medford, MA.

Duque, J. C., L. Anselin, and S. J. Rey. 2012. The max-p-regions problem. *Journal of Regional Science* 52 (3):397–419. doi: 10.1111/j.1467-9787.2011.00743.x.

Erwig, M., and M. Schneider. 1997. Vague regions. In *International symposium on spatial databases*, ed. M. Scholl and A. Voisard, 298–320. Berlin: Springer.

Fonte, C. C., and W. A. Lodwick. 2004. Areas of fuzzy geographical entities. *International Journal of Geographical Information Science* 18 (2):127–50. doi: 10.1080/13658810310001620933.

Forest, B. 2004. Information sovereignty and GIS: The evolution of "communities of interest" in political redistricting. *Political Geography* 23 (4):425–51. doi: 10.1016/j.polgeo.2003.12.010.

Gao, S., K. Janowicz, D. R. Montello, Y. Hu, J.-A. Yang, G. McKenzie, Y. Ju, L. Gong, B. Adams, and B. Yan. 2017. A data-synthesis-driven method for detecting and extracting vague cognitive regions. *International Journal of Geographical Information Science* 31 (6):1–127. doi: 10.1080/13658816.2016.1273357.

Gao, S., Y. Liu, Y. Wang, and X. Ma. 2013. Discovering spatial interaction communities from mobile phone data. *Transactions in GIS* 17 (3):463–81. doi: 10.1111/tgis.12042.

Good, P. I. 2000. *Permutation tests: A practical guide to resampling methods for testing hypotheses*. New York: Springer Science & Business Media.

Goyal, M. K., and A. Sharma. 2016. A fuzzy c-means approach to regionalization for analysis of meteorological droughts in western India. *Natural Hazards* 84 (3):1831–47. doi: 10.1007/s11069-016-2520-9.

Guo, D. 2008. Regionalization with dynamically constrained agglomerative clustering and partitioning (REDCAP). *International Journal of Geographical Information Science* 22 (7)801–23.

Guo, Q., Y. Liu, and J. Wieczorek. 2008. Georeferencing locality descriptions and computing associated uncertainty using a probabilistic approach. *International Journal of Geographical Information Science* 22 (10):1067–90. doi: 10.1080/13658810701851420.

Hall, O., and W. Arnberg. 2002. A method for landscape regionalization based on fuzzy membership signatures. *Landscape and Urban Planning* 59 (4):227–40. doi: 10.1016/S0169-2046(02)00050-6.

Helbich, M., W. Brunauer, J. Hagenauer, and M. Leitner. 2013. Data-driven regionalization of housing markets. *Annals of the Association of American Geographers* 103 (4):871–89. doi: 10.1080/00045608.2012.707587.

Herschlag, G., S. Dibaeinia, and E. Bonica. 2022. Simulated redistricting plans for the analysis and evaluation of redistricting. *arXiv preprint*.

Holtz, D., M. Zhao, S. G. Benzell, C. Y. Cao, M. A. Rahimian, J. Yang, J. Allen, A. Collis, A. Moehring, T. Sowrirajan, et al. 2020. Interdependence and the cost of uncoordinated responses to Covid-19. *Proceedings of the National Academy of Sciences of the United States of America* 117 (33):19837–43. doi: 10.1073/pnas.2009522117.

Hwang, S., and J.-C. Thill. 2009. Delineating urban housing submarkets with fuzzy clustering. *Environment and Planning B* 36 (5):865–82. doi: 10.1068/b34111t.

James, P. E. 1952. Toward a further understanding of the regional concept. *Annals of the Association of American Geographers* 42 (3):195–222. doi: 10.1080/00045605209352091.

Jin, M., L. Gong, Y. Cao, P. Zhang, Y. Gong, and Y. Liu. 2021. Identifying borders of activity spaces and quantifying border effects on intra-urban travel through spatial interaction network. *Computers, Environment and Urban Systems* 87:101625. doi: 10.1016/j.compenvurbsys.2021.101625.

Kang, Y., S. Gao, Y. Liang, M. Li, J. Rao, and J. Kruse. 2020. Multiscale dynamic human mobility flow dataset in the US during the Covid-19 epidemic. *Scientific Data* 7 (1):390. doi: 10.1038/s41597-020-00734-5.

Kruse, J., S. Gao, Y. Ji, K. Levin, Q. Huang, and K. R. Mayer. 2024. Identifying rich clubs in spatiotemporal interaction networks. *Annals of the American Association of Geographers* 115 (4):899–922. doi: 10.1080/24694452.2025.2464806.

Kruse, J., S. Gao, Y. Ji, D. P. Szabo, and K. R. Mayer. 2024. Bringing spatial interaction measures into multi-criteria assessment of redistricting plans using interactive web mapping. *Cartography and Geographic Information Science* 51 (4):513–32. doi: 10.1080/15230406.2023.2264750.

Leung, Y. 1999. Fuzzy sets approach to spatial analysis. In *Practical applications of fuzzy technologies*, 267–300. New York: Springer.

Li, B., S. Gao, Y. Liang, Y. Kang, T. Prestby, Y. Gao, and R. Xiao 2020. Estimation of regional economic development indicator from transportation network analytics. *Scientific Reports* 10 (1):1–15.

Li, M., S. Gao, F. Lu, K. Liu, H. Zhang, and W. Tu. 2021. Prediction of human activity intensity using the interactions in physical and social spaces through graph convolutional networks. *International Journal of Geographical Information Science* 35 (12):2489–516. doi: 10.1080/13658816.2021.1912347.

Liang, Y., J. Zhu, W. Ye, and S. Gao. 2022. Region2vec: Community detection on spatial networks using graph embedding with node attributes and spatial interactions. In *Proceedings of the 30th ACM SIGSPATIAL Conference*, ed. M. Renz and M. Sarwat, 1–4. New York: ACM.

Liang, Y., J. Zhu, W. Ye, and S. Gao. 2025. GeoAI-enhanced community detection on spatial networks with graph deep learning. *Computers, Environment and Urban Systems* 117:102228. doi: 10.1016/j.compenvurbsys.2024.102228.

Liao, C., D. Brown, D. Fei, X. Long, D. Chen, and S. Che. 2018. Big data-enabled social sensing in spatial analysis: Potentials and pitfalls. *Transactions in GIS* 22 (6):1351–71. doi: 10.1111/tgis.12483.

Liu, Y., X. Liu, S. Gao, L. Gong, C. Kang, Y. Zhi, G. Chi, and L. Shi. 2015. Social sensing: A new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers* 105 (3):512–30. doi: 10.1080/00045608.2015.1018773.

Liu, Y., Y. Yuan, and S. Gao. 2019. Modeling the vagueness of areal geographic objects: A categorization system. *ISPRS International Journal of Geo-Information* 8 (7):306. doi: 10.3390/ijgi8070306.

Luo, P., and D. Zhu. 2022. Sensing overlapping geospatial communities from human movements using graph affiliation generation models. In *Proceedings of the 5th ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery*, ed. D. Lunga and S. Newsam, 1–9. New York: ACM.

Makse, T. 2012. Defining communities of interest in redistricting through initiative voting. *Election Law Journal* 11 (4):503–17. doi: 10.1089/elj.2011.0144.

Malone, S. J. 1997. Recognizing communities of interest in a legislative apportionment plan. *Virginia Law Review* 83 (2):461. doi: 10.2307/1073783.

McBratney, A. B., and I. O. Odeh. 1997. Application of fuzzy sets in soil science: Fuzzy logic, fuzzy measurements and fuzzy decisions. *Geoderma* 77 (2–4):85–113. doi: 10.1016/S0016-7061(97)00017-7.

McKenzie, G., and B. Adams. 2017. Juxtaposing thematic regions derived from spatial and platial user-generated content. In *13th International Conference on Spatial Information Theory (COSIT 2017)*, ed. E. Clementini, 20:1–20:14. LAquila, Italy: Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

Montello, D. R., A. Friedman, and D. W. Phillips. 2014. Vague cognitive regions in geography and geographic information science. *International Journal of Geographical Information Science* 28 (9):1802–20. doi: 10.1080/13658816.2014.900178.

Montello, D. R., M. F. Goodchild, J. Gottsegen, and P. Fohl. 2003. Where's downtown?: Behavioral methods for determining referents of vague spatial queries. *Spatial Cognition & Computation* 3 (2):185–204. doi: 10.1207/S15427633SCC032&3_06.

Mu, L., F. Wang, V. W. Chen, and X.-C. Wu. 2014. A place-oriented, mixed-level regionalization method for constructing geographic areas in health data dissemination and analysis. *Annals of the Association of American Geographers* 105 (1):48–66. doi: 10.1080/00045608.2014.968910.

Murphy, A. B. 2004. Territorial ideology and interstate conflict. In *The geography of war and peace: From death camps to diplomats*, ed. C. Flint, 280–96. London: Oxford Academic.

National Conference of State Legislatures. 2021a. Redistricting criteria. Accessed April 2025. https://www.ncsl.org/redistricting/redistricting-criteria.

National Conference of State Legislatures. 2021b. Redistricting law 2020. Accessed April 2025. https://www.ncsl.org/research/redistricting/redistricting-law-2020.aspx.

National Conference of State Legislatures. 2021c. State restrictions on partisan gerrymandering. Accessed April 2025. https://www.ncsl.org/redistricting/partisan-gerrymandering-and-the-courts.

Nelson, G. D., and A. Rae. 2016. An economic geography of the United States: From commutes to megaregions. *PLoS ONE* 11 (11):e0166083. doi: 10.1371/journal.pone.0166083.

Pal, N. R., and J. C. Bezdek. 1995. On cluster validity for the fuzzy *c*-means model. *IEEE Transactions on Fuzzy Systems* 3 (3):370–79. doi: 10.1109/91.413225.

Ramachandran, G., and D. Gold. 2018. Using outlier analysis to detect partisan gerrymanders: A survey of current approaches and future directions. *Election Law Journal: Rules, Politics, and Policy* 17 (4):286–301. doi: 10.1089/elj.2018.0503.

Rao, J., S. Gao, M. Miller, and A. Morales. 2022. Measuring network resilience via geospatial knowledge graph: A case study of the us multi-commodity flow network. In *Proceedings of the 1st ACM*

SIGSPATIAL International Workshop on Geospatial Knowledge Graphs, ed. M. Renz and M. Sarwat, 17–25. New York: ACM.

Ratti, C., S. Sobolevsky, F. Calabrese, C. Andris, J. Reades, M. Martino, R. Claxton, and S. H. Strogatz. 2010. Redrawing the map of Great Britain from a network of human interactions. *PLoS ONE* 5 (12):e14248. doi: 10.1371/journal.pone.0014248.

Schultz, K. A. 2017. Mapping interstate territorial conflict: A new data set and applications. *Journal of Conflict Resolution* 61 (7):1565–90. doi: 10.1177/0022002715620470.

Shaw, S.-L., and D. Sui. 2020. Understanding the new human dynamics in smart spaces and places: Towards a splatial framework. *Annals of the American Association of Geographers* 110 (2):339–48. doi: 10.1080/24694452.2019.1631145.

Shaw, S.-L., and H. Yu. 2009. A GIS-based time-geographic approach of studying individual activities and interactions in a hybrid physical–virtual space. *Journal of Transport Geography* 17 (2):141–49. doi: 10.1016/j.jtrangeo.2008.11.012.

Shi, W., and K. Liu. 2004. Modeling fuzzy topological relations between uncertain objects in a GIS. *Photogrammetric Engineering & Remote Sensing* 70 (8):921–29. doi: 10.14358/PERS.70.8.921.

Smith, B., and D. M. Mark. 2003. Do mountains exist? Towards an ontology of landforms. *Environment and Planning B* 30 (3):411–27. doi: 10.1068/b12821.

Stephanopoulos, N. O. 2012a. Redistricting and the territorial community. *University of Pennsylvania Law Review* 1:1379–477.

Stephanopoulos, N. O. 2012b. Spatial diversity. *Harvard Law Review* 1:1903–2010.

Storper, M. 1997. *The regional world: Territorial development in a global economy*. New York: Guilford.

Taherdoost, H., and M. Madanchian. 2023. Multi-criteria decision making (MCDM) methods and concepts. *Encyclopedia* 3 (1):77–87. doi: 10.3390/encyclopedia3010006.

Tobler, W. R. 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46 (Suppl. 1):234–40. doi: 10.2307/143141.

Van Ranst, E., H. Tang, R. Groenemam, and S. Sinthurahat. 1996. Application of fuzzy logic to land suitability for rubber production in peninsular thailand. *Geoderma* 70 (1):1–19. doi: 10.1016/0016-7061(95)00061-5.

Varzi, A. C. 2001. Vagueness in geography. *Philosophy & Geography* 4 (1):49–65. doi: 10.1080/10903770124125.

Wang, F., and G. B. Hall. 1996. Fuzzy representation of geographical boundaries in GIS. *International Journal of Geographical Information Systems* 10 (5):573–90. doi: 10.1080/02693799608902098.

Webster, G. R. 2013. Reflections on current criteria to evaluate redistricting plans. *Political Geography* 32:3–14. doi: 10.1016/j.polgeo.2012.10.004.

Xie, J., B. K. S. Kelley, and B. K. Szymanski. 2013. Overlapping community detection in networks: The state-of-the-art and comparative study. *ACM Computing Surveys* 45 (4):1–35. doi: 10.1145/2501654.2501657.

Xu, Y., D. Zou, S. Park, Q. Li, S. Zhou, and X. Li. 2022. Understanding the movement predictability of international travelers using a nationwide mobile phone dataset collected in South Korea. *Computers, Environment and Urban Systems* 92:101753. doi: 10.1016/j.compenvurbsys.2021.101753.

Yang, J., and J. Leskovec. 2013. Overlapping community detection at scale: A nonnegative matrix factorization approach. In *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining*, ed. M. Renz and M. Sarwat, 587–96. New York: ACM.

Yang, J., J. McAuley, and J. Leskovec. 2013. Community detection in networks with node attributes. In *2013 IEEE 13th International Conference on Data Mining*, ed. P. S. Yu, 1151–56. Piscataway, NJ: IEEE.

Yuan, J., Y. Zheng, and X. Xie. 2012. Discovering regions of different functions in a city using human mobility and pois. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ed. M. Renz and M. Sarwat, 186–94. New York: ACM.

Zadeh, L. A. 1965. Fuzzy sets. *Information and Control* 8 (3):338–53. doi: 10.1016/S0019-9958(65)90241-X.

Zhang, H., X. Zhou, Y. Yang, H. Wang, X. Ye, and G. Tang. 2024. Advancing process-oriented geographical regionalization model. *Annals of the American Association of Geographers* 114 (10):2388–413. doi: 10.1080/24694452.2024.2380893.

Zhu, D., F. Zhang, S. Wang, Y. Wang, X. Cheng, Z. Huang, and Y. Liu. 2020. Understanding place characteristics in geographic contexts through graph convolutional neural networks. *Annals of the American Association of Geographers* 110 (2):408–20. doi: 10.1080/24694452.2019.1694403.

JACOB KRUSE is a PhD Candidate in the Department of Geography at the University of Wisconsin–Madison, WI 53706. E-mail: jikruse@wisc.edu. His research interests include redistricting, public health, geospatial data science, and GeoAI.

SONG GAO is an Associate Professor of Geographic Information Science in the Department of Geography, University of Wisconsin–Madison, WI 53706. E-mail: song.gao@wisc.edu. His research interests include geospatial data science, GeoAI, social sensing, spatial networks, and human mobility.

KENNETH R. MAYER is a Professor in the Department of Political Science, University of Wisconsin–Madison, WI 53706. E-mail: krmayer@wisc.edu. His research interests include redistricting and American politics.