Position Paper

MICHAEL GOODCHILD
Department of Geography
University of California, Santa Barbara
Email: good@geog.ucsb.edu

The position papers submitted for the meeting raise a fascinating array of issues, ranging from representation and analysis to the social context. I would like to focus on two issues that seem to me not to have received much attention in the papers, but to be important in addressing the topic of the meeting.

Inference from spatiotemporal pattern

Geographers have long been interested in phenomena distributed in space and time, and in making inferences from pattern to process. Every dynamic process leaves a characteristic footprint or pattern on the Earth's surface, from which it is possible to infer process, though subject to the problem of equifinality, or the tendency of multiple processes to produce identical patterns. Some of the most compelling examples concern disease, and inferences about the causal mechanisms that have been derived from the spatial patterns of disease morbidity. Many of these studies have been carried out without explicit respect to time, either because data were available only in cross-sectional form, or because of the difficulties of visualizing and making inferences from truly spatiotemporal data. This situation is changing rapidly, however, because of the rapid growth in new spatiotemporal data sources and tools. Time adds much useful information and allows us to come closer to inferences about cause.

Applied to the specific topic of this specialist meeting, this logic implies that one can study spatiotemporal patterns to infer the existence and operation of social networks, and to build models of their structure. Applied to the data collected by Tilley on the Swing Rebellion of 1832 in England, for example, it suggests that we can learn about the structure of social networks in this pre-Internet age by observing the spatiotemporal diffusion of the rebellion's events. Geographers might distinguish in this example between spatial diffusion, in which news travels from person to person through physical interaction, and hierarchical diffusion, in which the contemporary print media were able to reach large numbers of people roughly simultaneously. The spatiotemporal pattern suggests that diffusion was initially spatial and then hierarchical.

Proximity provides the key to distinguishing spatial from hierarchical diffusion, since under spatial diffusion an event must be close to at least one prior event. Proximity can be modeled using a variety of distance-decay functions, including the negative exponential, which unlike negative powers has the advantage of being both well-behaved and supported by sound theory. Hierarchical diffusion can also be expected to display effects of proximity, though in this case the relevant distance is that between the producer and the consumer of media. Electronic communication, starting with the telegraph in the mid 1800s, has the potential to remove the effects of proximity entirely, driving the parameter of the negative exponential function to zero. Nevertheless there is abundant evidence that proximity effects persist in many instances of electronic communication, for a variety of reasons.

While the literature includes many instances of both hierarchical and spatial diffusion, I am not aware of any attempts to combine them in a single model, though this seems essential if we are to theorize about diffusion in an Internet-based world. One way to do this would be by mixing two pdfs, one the negative exponential of traditional spatial diffusion, and the other the uniform distribution implied by perfect communication. It would be interesting to do this within the context of a spatial interaction model, where the origin and destination mass terms can be included. Another approach would be through a small-world network model, which includes both forms of interaction.

Digital representation of embedded networks

The analysis and modeling of phenomena embedded in space and time has received an enormous boost in the past four decades. GPS, remote sensing, and GIS combine to provide an abundance of data, many of them born digital, and allow easy manipulation and sharing. A GIS today is capable of virtually any systematic operation on spatial data, and has been described as a new medium of communicating what is known about the surface and near-surface of the planet. However almost all attention in the development of GIS has been directed to maps, and the map remains a powerful metaphor for the contents of a GIS database. Maps are very efficient ways of showing what exists at each point in a mapped section of the Earth's surface, and humans are adept at connecting the dots to infer the existence of lines and areas, and to associate annotation with these features. However maps are not ideal media for communicating spatiotemporal information, since maps are inherently static once printed. Video addresses this problem, but is similarly limited in being able to display only one color at any point at any time.

Information about social networks embedded in space and time falls into the category of *binary* geographic information, or information about *pairs* of locations. Connections, interactions, flows, and movements all are of this nature, replacing the unary form $<\mathbf{x},z>$ with the binary form $<\mathbf{x}_1,\mathbf{x}_2,z>$. In the object-oriented paradigm that now dominates the design of GIS databases, such information would typically appear in association classes, since it qualifies the relationship between two features rather than the properties of a single feature. Visualization of binary geographic information remains a hard problem. We see this in the "fish-tank" views that assign the third (vertical) dimension to time in displaying large numbers of tracks, creating a confusing mass of lines that are difficult to interpret. New methods of measuring the similarity of tracks, and clustering them, may provide some clarity, but the appropriate metrics of similarity may be hard to define.