Time geography and space–time prism

Harvey J. Miller  
*The Ohio State University, USA*

Time geography is a constraints-oriented approach to understanding human activities in space and time. Time geography recognizes that humans have fundamental spatial and temporal limitations: people can physically only be in one place at a time and activities occur at a sparse set of places for limited durations. Participating in an activity requires allocating scarce available time to access and conduct the activity. Constraints on activity participation include the location and timing of anchors that compel presence (such as home and work), the time budget for access and activity, and the ability to trade time for space in using mobility or information and communication technologies (ICTs). Time geography is a physical not a behavioral theory; it highlights the necessary spatiotemporal conditions for human activities, but does not explain the sufficient events that lead to specific activities. But since these necessary conditions vary by individual and situation, time geography supports an approach to understanding human and environmental systems that recognizes individual histories and the importance of geographic context.

Time geography is consistent with some core ideas in fields such as geography, transportation, urban science, social sciences, and environmental sciences. These include an integrated perspective on human and physical phenomena, the need to build macro-level explanations from micro-level processing, and situating human activities within context. Basic time geographic concepts, such as events being sparsely distributed in time and space, limited time availability, and trading time for space to access activities, seem mundane, since they are common and correspond with everyday experience. But this is why time geography is needed: these seemingly banal but utterly crucial factors in our scientific explanations of human behavior should not be neglected. Time geography provides a framework that demands recognition of the fundamental constraints underlying human experience and also provides an effective conceptual system for keeping track of these conditions.

Time geography originates from Professor Torsten Hägerstrand (1916–2004), a Swedish geographer who spent his career at the University of Lund. He nurtured the ideas for a long time, but time geography emerged dramatically to the international scientific community with a now-famous 1969 presidential address to the Regional Science Association (Hägerstrand 1970). Hägerstrand was concerned that human geography and regional science were neglecting much of what comprises a livable world. He also wanted to provide a counterbalance to increasing specialization and fragmentation in science, technology, and administration by offering a more holistic view of human activities. Hägerstrand believed in the integrative power of the regional approach in geography, but felt that it needed to be more inclusive and rich. Time geography upgrades the regional approach by describing the “bare skeleton” of spatiotemporal conditions and constraints that emerge from the interplay of historical and geographical factors.

Time geography is an active and flourishing research domain one half-century after its initial
conceptualization in the 1960s. This ranks time geography among elite and enduring scientific ideas. Time geography has endured for good reasons. It is beautiful: the geometry of the space–time path and prism (discussed below) and their relationships are intuitive and appealing. It is elegant: it can help to explain much with few basic principles. It is robust: it can help us understand a wide range of phenomena in human and linked human-environmental systems. It is sensitive: it treats people as individuals and recognizes social differences across a wide range of factors (gender, age, socioeconomic status, culture) as well as geographic context. It is ecological: it connects the individual to the aggregate, balancing nomothetic law seeking with context and situation, as well as balancing agency and structure. Finally, it is practical: Hägerstrand was prescient when thinking about how to keep track of the basic spatial and temporal existential facts of human activities. It is now possible to collect, store, manage, and analyze individual-level data on mobile objects and human activities with ease and power that would seem magical from the perspective of the 1960s.

**Classical time geography**

Time geography recognizes three major types of constraints on human activities. Capability constraints limit the activities of individuals through their own physical capabilities and/or available resources. People need to conduct maintenance activities such as eating and sleeping; these require time and place. Also, individuals with private automobiles can generally travel faster than individuals who walk or rely on public transportation. Coupling constraints define where, when, and for how long an individual has to join with other individuals for shared activities such as work, meetings, and classes. Authority constraints are fiat restrictions over particular space–time domains. For example, a shopping mall or gated community can make it difficult and illegal to enter at designated times, while a public street cannot.

**Space–time path**

Activities such as personal and domestic maintenance, work, shopping, health care, education, and recreation are sparsely distributed in time and space; they are available for limited duration at relatively few locations. Participating in activities requires trading time for space to access these locations at their available times. The space–time path highlights these requirements. Figure 1 illustrates a space–time path between activity stations. Stations are places where activities can occur; classical time geography treats these as tubes designating their locations in space and availability in time (e.g., work hours, operating hours for a store, appointments, scheduled lectures). The classical space–time path focuses on physical mobility and interaction. Virtual interaction via ICTs is possible; for example, a telephone call can be represented as a connection between two space–time paths. However, virtual interaction is muted relative to physical interaction in the classic theory.

Time geography also classifies activities based on their flexibility for an individual. Fixed activities are those that cannot be easily rescheduled or relocated (e.g., work, meetings), while flexible activities can be more easily rescheduled and/or occur at more than one location (e.g., shopping, recreation). These categories can be arbitrary; for example, a software developer can code in an office or a cafe. Nevertheless, the dichotomy provides an effective means for understanding how the location and timing of some activities condition accessibility to other activities. Fixed activities act as space–time anchors because other
Space–time prism

The *space–time prism* (STP) highlights the influence of space–time anchors on the ability to participate in flexible activities. The STP is the envelope of all possible space–time paths between known locations and times. Figure 2 illustrates a planar STP. In this case, two *space–time anchors* frame a prism. Anchors correspond to known locations and times for the mobile object. These are often (but not always) the locations and times of fixed activities that compel presence. A maximum travel speed represents the object’s mobility capabilities; in classic time geography this is a uniform across space and time. (Time geography uses the term “velocity,” but this is incorrect since velocity implies magnitude and direction. “Speed” implies magnitude only and is, therefore, the more appropriate term.) Given these anchors and the speed limit, the prism defines the envelope all possible space–time paths between the anchors. The spatial footprint of the STP is the *potential path area* (PPA); this is the region in space that is accessible to the moving object.

The prism in Figure 2 is general since it accounts for stationary activity time and has two anchors that are spatially separate. A prism without stationary activity time consists only of its *time-forward cone* rooted at the first anchor and its *time-backward cone* rooted at the second anchor. The STP will have a larger volume and PPA, as no stationary activity means more time to be mobile. The two anchors may also be coincident spatially; this prism consists of two right cones instead of oblique cones as in Figure 2. A prism may also have only one anchor. An STP with its first anchor only is a *time-forward cone* delimiting all destinations that can be reached from that origin within a specific time limit. Conversely, a prism with its second anchor only is a *time-backward cone* delimiting all origins that can reach that destination within a specific time limit.

The STP measures *accessibility*: the ability for an individual to travel and participate in activities and the amount of time available for activity participation at locations. An activity at a station is not feasible unless that station intersects with the prism spatially and temporally, the latter for at least as long as the minimum activity time required. This delimits the subset of opportunities in an environment that is available to a person based on their STP constraints.

The STP provides a dramatically different image of accessibility than place-based accessibility measures, such as those based on spatial interaction (“gravity”) models or simply counting the number of opportunities near a person’s...
home and work location. The locations and timings of fixed activities vary by age/life cycle stage, socioeconomic status and culture. For example, STP-based accessibility measures capture gender differences in accessibility due to household organization that are missed by place-based accessibility measures such as home–work commute length. Place-based accessibility measures assume everyone at a place, such as home and work, has the same space–time scheduling constraints, while people-based measures derived from the space–time prims recognize individual capability constraints that are masked through homogenization by place. Time geography also facilitates understanding how the temporal organization of service and trading hours can have differential impacts beyond the locations of services and businesses.

**Bundling and intersections**

Time geography recognizes two types of relationships between paths and prisms. **Bundling** refers to the convergence in space and time of

![A planar space–time prism.](image)

**Figure 2** A planar space–time prism.
two or more paths. Path bundling is necessary (although not sufficient) evidence of shared activities and individuals meshing their space–time activities to participate in projects. Bundling can occur when objects are in motion or stationary; examples of the former include public transportation and ride-sharing. Intersection is the condition of two or more time geographic features sharing some locations in space with respect to time. Two or more people cannot physically meet unless their STPs intersect, or a path is within an STP.

Bundling and intersections are necessary conditions for the emergence of broader space–time activity systems, such as a university or a city. Time geography is an ecological theory; it considers interactions between individuals and between the individual and the aggregate. Bundling and intersections require individuals to synchronize (coordinate over time) as well as synchronize (coordinate over space). Coordination also occurs at multiple scales. Individuals must conduct projects consisting of sequenced activities linked by mobility events to meet larger goals such as hosting a dinner party. Cities and regions are time systems balancing the supply of available time and demands on that time. This is a rich and intricate conceptualization of cities and societies; Alan Pred memorably referred to a “ballet of adjustments” as disruptions propagate through a time system due to activities and projects changing to meet the new space–time requirements for participation and interaction.

**Analytical time geography**

Classical time geography is conceptually rich but limited analytically. The rise of geographic information systems (GIS) motivated renewed interest in time geography, particularly in relaxing strict assumptions such as the maximum speed constraining an STP being uniform in space and time. The development and deployment of location-aware technologies (LATs), such as the global positioning system (GPS), mobile phones, and radiofrequency identification (RFID) chips, have greatly expanded capabilities for collecting data on mobile objects and have led to the development of mobile objects databases and mobility mining or exploratory analysis of mobile objects data. These scientific and technological developments led to the development of analytical time geography and supporting GIS software.

**Path analytics**

Location-aware technologies do not generate space–time paths directly; they generate a temporal sequence of spatial locations that are used to construct the path. There are several ways to generate this temporal sequence. Event-based recording captures the time and location when a specified event occurs; for example, a person texting or calling using a mobile phone. Time-based recording captures mobile object positions at regular time intervals; this typifies GPS receivers. In change-based recording a capture occurs when the position of the object is sufficiently different from a previous location; this includes dead-reckoning methods as well as some mobile objects data technologies that avoid recording locations to manage data volume. Location-based recording occur when a mobile object comes close to locations where sensors are located; examples include stationary radiofrequency identification and Bluetooth sensors.

The simplest and most common way to generate a path is linear interpolation: assume the object followed the straight-line segment between recorded locations. This works well for time-based and change-based recording with high capture frequencies. Event-based recording creates more issues since events are often not
very frequent, and the spatial resolution can be coarse and variable; this is often the case with mobile phone data resolved only to membership in a service cell. Even coarser are paths derived from location-based recording; these are often simply sequences of visits at the sensor locations. Space–time paths can also be matched to other spatially referenced data such as transportation networks.

The ease of collecting space–time path data from location-aware technologies often comes at the expense of path semantics: details about the moving objects such as the reasons for mobility behavior. Semantics can be recovered by overlaying paths with other georeferenced data. This method can produce errors related to data inaccuracies and intrinsic ambiguities. For example, GPS positional error means that it can be hard to tell if a person is inside or outside of an activity location, for example, a coffee house. If the person is inside the coffee house, what was she doing – dining, working, socializing, or some combination of the above? Often, there is no unambiguous link between locations and activities, especially in an era with near-ubiquitous access to information and communication. Methods for recovering path semantics include decomposing the trajectory into a sequence of moves and stops, and annotating these sequences based on map matching with background geographic information. Also available are advanced data mining techniques such as machine learning algorithms.

Table 1 summarizes the fundamental measurements available from space–time paths. The primitive position allows the derivation of distance, direction, and spatial extent in the object’s movement; an example is spatial range methods. From these primary derivatives the spatial distribution of the object, changes in direction, and the path’s sinuosity can be derived. From the temporal primitives instant and interval the duration of the object’s movement and its speed, as well as the temporal distribution of the object’s movement and changes in duration, can be derived. Primary derivatives from combining spatial and temporal intervals include the object’s speed (a scalar) and velocity (a vector), and corresponding secondary derivatives acceleration and an approaching rate with respect to a location, boundary or region.

### Prism analytics

It is difficult to analytically describe the entire STP. However, it is easy to describe its spatial extent at a moment in time; this can serve as

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Primitive</th>
<th>Derivatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial</td>
<td>Position</td>
<td>Distance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Duration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Acceleration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spatial distribution</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sinuosity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Change of direction</td>
</tr>
<tr>
<td>Temporal</td>
<td>Instance</td>
<td>Duration</td>
</tr>
<tr>
<td></td>
<td>Interval</td>
<td>Travel time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temporal distribution</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Change of duration</td>
</tr>
<tr>
<td>Spatiotemporal</td>
<td>Speed</td>
<td>Acceleration</td>
</tr>
<tr>
<td></td>
<td>Velocity</td>
<td>Approaching rate</td>
</tr>
</tbody>
</table>

the basis for a wide range of prism analytics. STP parameters are \( \{x_i, x_j, t_i, t_j, s_{ij}, a_{ij}\} \) where \( x_i, x_j \) are the first and second anchor locations with associated departure and arrival times \( t_i, t_j \) respectively, \( s_{ij} \) is the maximum travel speed, and \( a_{ij} \) is the stationary activity time. At a moment in time \( t \in (t_i, t_j) \), the spatial extent of an STP (denoted by \( Z_{ij}(t) \)) is the intersection of three convex spatial sets: (i) the future disc \( f_i(t) \) comprising all locations that can be reached from the first anchor by time \( t_i + t \); (ii) the past disc \( p_j(t) \) encompassing all locations at time \( t \) that can reach the second anchor by time \( t_j - t \); and (iii) the geo-ellipse \( g_{ij} \) that constrains the prism locations to account for any stationary activity time:

\[
Z_{ij}(t) = \{ f_i(t) \cap p_j(t) \cap g_{ij}\} \\
f_i(t) = \{ x \mid \| x - x_i \| \leq (t - t_i)s_{ij} \} \\
p_j(t) = \{ x \mid \| x - x_j \| \leq (t_j - t)s_{ij} \} \\
g_{ij} = \{ x \mid \| x - x_i \| + \| x - x_j \| \leq (t_j - t_i - a_{ij})s_{ij} \}
\]

Figure 3 provides an illustration. This definition of the STP is not limited to 2-D space. In 1-D space, the sets described by equations (2)–(4) are line segments. In 2-D space, the discs are circles and the geo-ellipse is an ellipse. In 3-D space, the discs are spheres and the geo-ellipse is a spheroid. There are scalable methods for calculating these objects and their intersections. Also, only one or two of the sets are relevant at any time, since the future and past disc change in size and may be enclosed by the other two objects. For example, with a general prism as in Figure 2, it is only necessary to solve (in the following order) a disc, a disc–ellipse intersection, an ellipse, a disc–ellipse intersection, and finally a disc. Similarly, finding path–prism intersections only requires testing if a point lies within a disc, ellipse, or a disc–ellipse intersection. Finding a prism–prism intersection requires solving for the intersection of two, three, or four of the sets based on the prisms’ morphologies at that moment in time. The worst case is a four-set intersection involving two discs and two ellipses (Miller 2005).

Over time, the future disc traces the time-forward cone with an apex at the first anchor, the past disc traces the time-backward with an apex at the second anchor and the geo-ellipse is a cylinder. Over time, the prism in space is a lens-like set that traces the PPA footprint. These dynamics can be computed at discrete moments in time using the static construction described above and reconstruct the corresponding spatiotemporal region for the prism.

Error and uncertainty in paths and prisms

Error in space–time paths results from two sources. Measurement error occurs when the recording of a mobile object’s location has noise or uncertainty. The left-hand side of Figure 4 illustrates this for one segment of a space–time path; the captured locations have some degree of spatial error. This is equivalent to the problem of error in line segments and polylines, a well-established topic in the GIS literature. Sampling error occurs when the captured locations undersample an object’s movement pattern. This creates a spatial uncertainty region equivalent to the STP over time. Therefore, a space–time path with sampling error can be viewed as a sequence of linked STPs. The right-hand side of Figure 4 illustrates this; the black curve is the actual path of the mobile object, the dashed red polyline indicates the interpolated movement, and the blue ellipses indicate the spatial uncertainty region surrounding each paired sample locations.
Combined measurement and sampling error in a space–time path is equivalent to an STP with measurement error. Therefore, if it is possible to solve the measurement error problem for STPs it is also possible to solve the problem of combined measurement and sampling errors in space–time paths. One strategy is Monte Carlo simulation, which involves generating multiple prism realizations based on sampling from the stochastic error distributions. Monte Carlo simulation is effective for theoretical investigations but it is cumbersome for applications: it is awkward to run prism error simulations when executing STP-based mobile objects database queries or measuring accessibility in a model.

Spatial error propagation theory provides an analytical strategy for analyzing prism error. Given a direct geographic measurement $\mathbf{x} = \mu_x + \varepsilon_x$, where $\mu_x$ is the true location vector and $\varepsilon_x$ is an error vector, it must be determined how that error propagates through a geographic operation $\mathbf{y} = f(\mathbf{x})$. If $f(\mathbf{x})$ is nonlinear, the error propagation is approximated using the first-order partial derivatives of the function. In the case of the analytical prism, the relevant functions are the intersections of the boundaries of spatial objects described by equations (2)–(4). These boundaries are:

- $f_i : \| \mathbf{x} - \mathbf{x}_i \| - (t - t_i)s_{ij} = 0$ (5)
- $p_j : \| \mathbf{x} - \mathbf{x}_j \| - (t_j - t)s_{ij} = 0$ (6)
- $g_{ij} : \| \mathbf{x} - \mathbf{x}_i \| + \| \mathbf{x} - \mathbf{x}_j \| - (t_j - t_i - a_{ij})s_{ij} = 0$ (7)

A problem is that equations (5)–(7) define these boundaries implicitly rather than explicitly (that is, in the form $f(x, y) = 0$ rather than $f(x) = y$), meaning that solving for the intersection points requires finding the roots of high order polynomials. This difficulty can be resolved using implicit function methods that allow calculation of the required first order partial derivatives without having their explicit functions. This allows the error propagation from STP parameters to the constructed STP and to STP intersections to be estimated analytically. However, some of the required circle and ellipse intersection cases are
still unsolved and the techniques are not tractable beyond two STP intersections. Still required are scalable approximations and heuristics for STP error based on these analytical techniques. The analytical techniques can also serve as bounds for improving the efficiency of simulation-based techniques.

The methods discussed above estimate error propagation from measured STP parameters to the constructed object. Another strategy is to construct objects that encompass the possible STP consistent with the prism parameters and their errors. A rough STP consists of the upper and lower bounds on an STP. A reliable STP is the set of space–time locations where an individual can conduct an activity and meet the second anchor constraints with a specified probability.

**Properties of the prism interior**

The STP is traditionally a binary concept; all locations within the prism interior are considered to be equally accessible. However, this is a simple characterization that masks intricate properties of the prism interior. Intuitively, it would be expected that the distribution of visit probabilities would be unequal: locations that are near the space–time axis connecting the prism anchors are more likely to be visited than locations near the prism boundary since there are more possible paths through the former.

One way to model visit probabilities within a planar STP is to use the theory of random walks (RWs). An RW is a stochastic process in discrete space and time; given a spatial lattice of discrete locations, an RW involves, at each time step, a random choice of direction. Results from RW theory confirm intuition that locations near a prism anchor in a forward-time or backward-time cone are more likely to be visited; visit probabilities follow a bivariate multinomial distribution centered on the prism anchor.

However, movement within an STP with two anchors is not completely random; the object needs to move from the first prism anchor to second anchor within the time budget. A directed random walk (DRW) is an RW process with directional biases. A DRW process constrained by an STP requires updating the RW direction biases at each step to account for the remaining time to travel to the second anchor given the speed limit. This suggests a visit probability distribution similar to a bivariate normal distribution that is centered on the axis connecting the anchors and moves along that axis with respect to time.

A technique for modeling directed movement in continuous time and space is Brownian bridges (BBs). A BB is a continuous stochastic process between two known values; it has been applied widely to model animal movement between recorded locations. While BB can capture directionally biased random movement between prism anchors, is it not constrained by a maximum speed and, therefore, does not capture the STP boundary; it is still possible for the object to travel outside the prism boundary. A truncated Brownian bridge (TBB) imposes a maximum speed on a BB process, fully representing STP constraints. Figure 5 illustrates visit probabilities at a moment in time within the PPA of a planar STP based on a TBB process. Figure 5 shows the two prism anchors with the spatial projection of the space–time axis connecting the anchors. The gray area is the PPA for the STP; these locations have zero probability of being occupied at this moment in time. The blue, yellow, orange and red colors show increasing probability that the locations may be occupied by the object at that moment in time. These results are also consistent with the assumption of a bivariate random distribution whose center moves with time along the axis connecting the
anchor points; however, rather than assumed it is derived from fundamental movement principles. Another strategy for estimating visit probabilities within planar STPs is to use spatial interpolation to infer unknown object locations from the known locations at the prism anchors. This assumes that the most likely locations are along the axis between the anchors, but as seen above this assumption is consistent with theory. Time-geographic density estimation (TDE) integrates kernel density estimation with the constraints imposed by an STP. This method can generate the fine-grained movement patterns of objects between anchors derived from sparse tracking data.

Other types of space–time prisms

A planar STP has interesting properties and is useful for applications where it can be assumed that objects move in constrained space with a uniform maximum speed. However, in some applications the assumptions of unconstrained space and uniform maximum speed are unrealistic; examples include vehicles within a transportation network or movement across terrain. In addition, the STP assumes unrealistic motions such as infinite rates of acceleration and deceleration.

A network time prism (NTP) is a prism subject to network routes and allowable speeds that vary by network arc and, possibly, time. Figure 6 illustrates an NTP for a network embedded in 2-D space with time along the third orthogonal dimension (Kuijpers and Othman 2009). The green subportion of the network arcs is the potential network area (PNA), the network analog of the planar PPA; these are the accessible locations within the network. The red polygon comprises the full NTP; this is the envelope of possible space–time paths between the anchors constrained by the network.

Calculating an NTP involves three steps. A precomputation step uses the planar PPA as an upper bound on the NPA; this speeds computation by eliminating network nodes that cannot possibly be in the NPA. It then solves for the shortest path tree in this subnetwork twice: once for travel from the first anchor and a second time for travel to the second anchor. The procedure assigns the earliest arrival time and latest departure time at each node. The second step builds the spatial footprint of the NTP by testing if only one of an arc’s end nodes is included in the NPA and (if so) solving for the NPA boundaries within the edge. The third step computes the full spatiotemporal region of the NTP based on the earliest arrival times and latest departure times at network vertices. These arrival and departure times dictate, for each node incident to an arc, whether the mobile object can traverse the full arc to an adjacent node or can only traverse partway and return to the original node. These cases correspond to the
rectangular and triangular regions in Figure 6, respectively. The computational bottleneck in the algorithm is the shortest path trees in the precomputation step; although this is only applied to a subportion of the network, this subnetwork can be large, as can the number of prisms to be processed. However, the Boolean query about whether two NTPs intersect can be solved analytically.

Similar to an STP, the parameters of an NTP can also be measured with uncertainty. An anchor region is the set of all possible locations and times for a prism anchor within a network arc. These regions are not necessarily continuous in space or time and can have nonuniform probabilities. There are two ways to calculate an NTP with anchor regions. The first way is to calculate the envelope of all NTPs having an anchor point within a given anchor region. The second way is to calculate, for any space–time point, the probability that an NTP with given anchor regions contains that point. As with the standard NTP, the computational bottleneck is the precomputation phase involving shortest path calculations.

A field time prism (FTP) is a planar STP where travel speeds (magnitude) or velocities (magnitude and direction) vary continuously across space. Fields are useful for describing movement across terrain or through water and air that are subject to currents. Fields can extend the NTP by treating edges as 2-D regions with continuously varying speeds or velocities due to factors such as traffic. In these cases, space–time paths have unobserved components corresponding to minimum cost curves through an inverse speed or velocity field rather than straight line segments through a uniform plane. The FTP generalizes the prism concept: the STP and the long standing concept of isochrones (curves of equal travel time) are special cases of the FTP. It also links time geography to the continuous transportation or urban fields tradition in quantitative geography and regional science.

Space–time prisms as described above are physically impossible; they assume that the mobile object can instantly accelerate and decelerate at locations such as the prism anchors and the prism’s sharp corners. Physical limitations on acceleration and deceleration mean that an STP is an overestimate of the true space–time region accessible to the object. While this may not matter for some STP applications, kinetics can make a difference for applications such as microscale movement (e.g., pedestrians, athletes), active transport modes such as bicycling, movement through media such as airplanes and ships, and animal behavior. It can also make a difference for applications in sustainable transportation, as vehicle fuel consumption and emissions are often dominated by acceleration events.

A kinetic time prism (KTP) is the set of all possible kinetic paths between two locations and times, where a kinetic path is a space–time path that obeys physical limits on acceleration and deceleration. Figure 7 illustrates a KTP in
Figure 7 A kinetic (solid line) and classic (dotted line) space–time prism in 1-D space.

1-D space and time, overlaid with a classic STP. Solving a KTP is scalable in 1-D space and time but complex in 2-D space. In 2-D space, only one-quarter of the prism must be solved; the remainder can be obtained through point and reflection symmetries. But calculating the first quarter of the KTP requires solving parametric functions describing the 1-D prism rotating around the line connecting the prism anchors. This can only be accomplished if it is assumed that the object’s initial heading is unknown; still open is the case where the object’s initial heading is known.

Path and prisms collections

With the proliferation of location-aware technologies generating mobile objects data, it is often the case that there are large collections of space–time paths and prisms to analyze. With these collections, it is often useful to summarize the paths or prisms through clustering (finding groups of similar paths or prism) or aggregation (forming a composite representative paths or prisms). It may also be desired to search through a collection of paths and prisms for other cases that resemble a reference path or prism. This requires methods for calculating path and prism similarity.

Path similarity is the degree of correspondence between two space–time paths. Geometric similarity measures focus only on the geometry, ignoring sequence and time; these include Euclidean and Hausdorff distances. Euclidean distances, such as the average, minimum and maximum distance are intuitive and scalable; however these measures are sensitive to noise and outliers. The Hausdorff distance is the maximum distance from a location in one path to the closest location in the other path. However, the Hausdorff distance can be misleading, as it does not take into account the temporal sequence of the path locations, only their geometry.

Dynamic similarity measures include Fréchet distance, dynamic time warping, longest common subsequences, and edit–distance functions. Unlike Hausdorff, the Fréchet distance takes into account the sequence of locations in the path and is, therefore, better suited to mobile objects. One way to think about the Fréchet distance is to imagine two space–time paths representing a person walking a dog: the Fréchet distance is the shortest leash that connects the two. Dynamic time warping measures the similarity between two sequences or trajectories based on the effort involved to stretch or compress time to get the sequences to match. Least common subsequence (LCSS) measures similarity based on the length of least common subsequence between two sequences. Edit-distance functions are a generalization of LCSS; these measure similarities between sequential patterns based on the cost of the insertion, deletion, and substitution operations required to transform one sequence into the other.

Other methods for analyzing collections of space–time paths include path clustering methods and spatial field methods. Path clustering methods find groupings of similar space–time
paths in collections of mobile objects; these are often based on path similarity measures such as the ones already discussed. *Spatial field methods* translate movement patterns of objects into fields or surfaces that summarize mobility and activity frequency by geographic location, allowing the identification of “hot spots” or locations with high mobility activity.

Similar to space–time paths, it may desired to analyze a collection of space–time prisms for patterns. This may involve clustering or aggregating prisms based on morphology, finding prisms similar to reference prism, and summarizing and aggregating prisms to discover synoptic patterns. However, to date there has been little research on these questions; it is a much less developed research topic than path comparison and summarization. A fundamental problem is to develop efficient numeric measures that summarize meaningful prism properties. It is easy to calculate basic properties such as prism volume and PPA size for classic prisms, as these have an elegant geometry consisting of cones and cylinders. More generally, measures are required that can capture a wide range of morphological properties. These properties may be straightforward for classic prisms, but more complex prisms such as NTPs, FTPs, and KTPs, or prisms with fuzzy or probabilistic boundaries have more complex shapes and, consequently, are more difficult to describe analytically.

One possible strategy for measuring planar space–time prism morphology is *shape analysis* or quantitative measures summarizing geometric form. Shape analytical techniques typically reduce geometric forms to a single numeric indicator or a small set of indicators; for example, roughness, perforation, and elongation. However, most shape measures are specific to 2-D space: required are scalable shape measures that can handle 3- or 4-D space (treating time as static) and the evolution of shape (treating time as dynamic).

A possible way to summarize NTP structure is through network indices based on graph theory. Network indices summarize network properties such as connectivity, the distribution of arc lengths/costs, and properties of shortest paths and tours within the network. Graph theory and network analysis is a very well established field that has been newly reinvigorated due to the rise of network science to analyze phenomena such as social network and cities. A research frontier is determining which existing network indices describe relevant properties of NTPs, and the development of new network indices tailored for NTPs.

**Joint accessibility and time ecology**

As mentioned, time geography recognizes bundling and intersections among paths and prisms as indicators of shared activities. The next step beyond calculating bundles and intersections are the higher-level properties associated with space–time coordination and dynamics.

*Collective motion methods* focus on individual object movement patterns within the context of a larger group of mobile objects. Distance-based measures search for collective patterns, such as flocks, by searching for moving objects that are densely connected in space given user-defined distance and time thresholds. Relative motion methods consider the individuals’ directions to detect collective patterns such as flocking, leadership and convergence in a group of moving objects. *Joint accessibility measures* use prism–prism intersections as a basis for analyzing the potential benefits to coordinated activity participation.

Static measures of collective motion and joint accessibility do not directly address the ultimate
goal of understanding the dynamic processes though which shared and interlinked activities form and intertwine in space and time. This highlights a weakness of time geography: its inability to specify an explanation of human activity that goes beyond the implications of space–time constraints limiting possible activities. Time geography can explain why infeasible activities did not occur, but it cannot explain why some feasible activities occurred while others did not. However, it would be difficult to build a dynamic, processes-based time ecology theory based only on the negative space implied by prisms and stations. A more promising way forward is to link time geography with modeling approaches such as activity-based travel demand models and agent-based models that naturally capture the emergent properties associated with individuals’ interactions over space and time. The natural fit between activity-based analysis and time geography is well recognized and well utilized in many models. However, the linkages between time geography and agent-based modeling are less developed, somewhat surprisingly given their complementarity.

**Time geography and virtual interaction**

Classical time geography recognizes the possibility of virtual interaction; for example, Hägerstrand (1970) shows a telephone call between two space–time paths. However, virtual interaction is neglected relative to travel and physical interaction in classical time geography. This is understandable given the era when time geography was initially formulated. But in the contemporary world virtual interaction is more pervasive and cannot be ignored when discussing physical mobility and activities.

Table 2 is a typology of possible communication modes based on space–time constraints.

**Synchronous presence** (SP) includes face-to-face conversations; this requires people involved to be physically present or proximal at the same time. **Asynchronous presence** (AP) requires co-location in space but not coincidence in time; examples include notes left on an office door or in a geocache. **Synchronous telepresence** (ST) requires coincidence in time but not in space; it includes media such as telephones, television, radio, and teleconferencing. **Asynchronous telepresence** (AT) does not require co-location in space or coincidence in time; it includes mail, e-mail, messages, and webpages. Classical time geography focuses on SP, recognizes but does not develop ST, and mostly ignores AP and AT.

There are at least two strategies for incorporating the wider range of communication modes into time geography. One strategy is to treat these as spatial and temporal relationships between space–time paths and prisms. SP requires the prism to intersect both spatially and temporally, while AP interactions only require spatial coincidence between the prisms (in other words, they share the same locations but at different times). ST requires temporal but not spatial coincidence

| **Table 2** Communication modes based on spatial and temporal constraints. |
|---|---|---|
| **Temporal Presence** | **Spatial Telepresence** |
| **Synchronous** | **SP** | **ST** |
| Face-to-face conversation | Telephone |
| Television |
| Radio |
| Teleconferencing |
| **Asynchronous** | **AP** | **AT** |
| Trail signs | Mail |
| Note left on office door | E-mail |
| Geocaches | Newspapers |
| Webpages |

TIME GEOGRAPHY AND SPACE–TIME PRISM

(the prisms share some interval in time but not locations in space), while AT only requires one prism to precede the other in time.

Another strategy for extending time geography is to make communication technology explicit using time geographic entities and derive space–time constraints based on those entities. A portal is a type of space–time station (Figure 1) where a person can access communication services. It includes a point location, and an access range and time intervals when the service is available. Portals correspond to real world entities such as wired Internet connections (a point location with zero range), wireless access points, and cellular telephone base stations (both characterized as point locations with different ranges). An individual can access a communication service only if his or her paths or prism intersect with the service footprint of an appropriate portal. When this occurs, it generates a message window or an interval of time when a message could be sent or received. Message windows can be compared using temporal predicates such as before, during, and after to determine which communication events are feasible. This approach is consistent with the perspective that temporal constraints dominate spatial constraints on telepresence.

Discovering time geographic knowledge

As the size of mobility datasets become richer and voluminous, the challenges move beyond developing scalable techniques for computing low-level time geographic properties and relationships to discovering higher-level time geographic knowledge. Mobility mining is a set of activities and techniques for discovering novel knowledge hidden in moving objects data. This involves three major activities: (i) trajectory reconstruction and management; (ii) space–time knowledge discovery; and (iii) space–time knowledge delivery.

Trajectory reconstruction and management requires processing the raw mobility data to obtain the space–time paths of the mobile objects as well as data structures and access methods for processing these paths efficiently. These include specialized data warehouse designs, mobile object indexing methods, and semantic trajectory compression methods for reducing storage requirements. Preserving locational privacy through security, anonymization, and other protocols is also a concern.

Space–time knowledge discovery involves processing the trajectories to discover patterns such as clusters and behavioral rules. As noted, path and prism similarity methods can be used in conjunction with clustering methods to provide synoptic summaries of large mobile objects databases. Collection motion methods can also scale to large mobile databases. Another technique is space–time association rule mining that discovers rules describing how objects move among a set of regions over time. Sequence mining techniques search for temporal patterns in mobile objects data. Periodic pattern mining search for recurrent patterns in sequential data includes finding locations that are repeatedly visited by mobile objects and the recurrent movement patterns between these locations. Under specific conditions, causal relationships can also be inferred from sequential pattern in movement data.

Space–time knowledge delivery from movement data requires methods for transparent management of the knowledge discovery process and linking discovered knowledge to real-world semantics. A strategy for managing mobility mining is through visual analytics. Visual analytics is an extension of scientific visualization, but rather than seeking insights into data, visual analytics seek insights into how data are processed during exploration and analysis. Visual
analytics for moving objects include techniques for analyzing and comparing entire trajectories, the distribution of properties within trajectories, synoptic visualization, and investigating movement within geographic context.

Connecting movement patterns to real-world movement semantics requires a conceptual system for describing low-level patterns and their relationships with higher-level behavior. System dimensions can include movement parameters (direct measures and derivatives; Table 1), the number of objects (one, group, or a cohort of similar objects such as people aged 20–30 years), path type (semantically continuous or discontinuous), influencing factors (the object’s intrinsic properties, spatial constraints on movement, environmental factors, and the influence of other agents), and scale/granularity of the data (spatial and temporal scales, temporal granularity). Generic patterns are low level and shared with many different types of mobile objects. These can be primitive (based on a single mobility parameter) or compound (based on multiple primitives and/or inter-object relations). Behavioral patterns are high level and specific to the object type. Examples include pursuit/evasion, fighting, flocking, and leader/follower. Unlike generic patterns, behavioral patterns are open-ended and can be added to the framework as more empirical movement patterns are discovered in different domains (Dodge, Weibel and Lautenschütz 2008).

SEE ALSO: Accessibility, in transportation planning; Agent–based modeling; Behavioral geography; Big data; Data model, moving objects; Geographic data mining; Geographic information science; Geographic information system; Geolocation services; GIS for transportation; Graph theory; Information and communications technology; Information technology and mobility; Network analysis; Regional geography; Representation: time; Representation: trajectories; Transport geography; Transport networks; Transport technology; Uncertainty

References


Further reading


