A Review of Quantitative Methods for Movement Data

Jed A. Long*, Trisalyn A. Nelson

Spatial Pattern Analysis & Research Lab, Department of Geography, University of Victoria, Victoria, BC, Canada

*corresponding author: jlong@uvic.ca

Pre-print of published version.

Reference:

DOI: http://dx.doi.org/10.1080/13658816.2012.682578

Disclaimer:
The PDF document is a copy of the final version of this manuscript that was subsequently accepted by the journal for publication. The paper has been through peer review, but it has not been subject to any additional copy-editing or journal specific formatting (so will look different from the final version of record, which may be accessed following the DOI above depending on your access situation).
Abstract

The collection, visualization, and analysis of movement data is at the forefront of geographic information science research. Movement data are generally collected by recording an object’s spatial location (e.g., XY coordinates) at discrete time intervals. Methods for extracting useful information, for example space-time patterns, from these increasingly large and detailed datasets have lagged behind the technology for generating them. In this article we review existing quantitative methods for analyzing movement data. The objective of this article is to provide a synthesis of the existing literature on quantitative analysis of movement data while identifying those techniques that have merit with novel datasets. Seven classes of methods are identified: 1) time geography, 2) path descriptors, 3) similarity indices, 4) pattern and cluster methods, 5) individual-group dynamics, 6) spatial field methods, and 7) spatial range methods. Challenges routinely faced in quantitative analysis of movement data include difficulties with handling space and time attributes together, representing time in GIS, and using classic statistical testing procedures with space-time movement data. Areas for future research include: investigating equivalent distance comparisons in space and time, measuring interactions between moving objects, development of predictive frameworks for movement data, integrating movement data with existing geographic layers, and incorporating theory from time geography into movement models. In conclusion, quantitative analysis of movement data is an active research area with tremendous opportunity for new developments and methods.
1 – Introduction

The study of movement in geographic information science (GISci) has followed a similar trajectory to the discipline of geography, whereby early work relied heavily on qualitative methods. In the 1960’s and 70’s the discipline of geography experienced a quantitative revolution whereby theory and methods were developed for explaining how place and space could be modeled as quantitative entities. The quantitative revolution produced developments in statistical methods designed specifically for spatial data, for instance spatial autocorrelation measures (Cliff and Ord 1973). Only later in the quantitative revolution did theoretical frameworks for quantitative analysis of movement emerge; most notably Hägerstrand’s (1970) time geography. As the quantitative revolution in geography sputtered in the late 1970’s (Johnston 1997) Hägerstrand’s ideas were primarily used as context for examining human behavior (e.g., Parkes and Thrift 1975, Pred 1981), rather than as an analytical toolkit for quantitative research. An exception is the work of Lenntorp (1976) and Burns (1979), which represent seminal pieces using time geography in quantitative analysis.

In the 1990’s, triggered by the development of geographic information systems (GIS), quantitative analysis again moved to the forefront of the geographic literature (Sheppard 2001). The term geographic information science (GISci) was coined to refer collectively to the science behind the collection, storage, representation, and analysis of geographic datasets (Goodchild 1992). The term amalgamated those interested in the study of geographic information including geographers, computer scientists, and statisticians. As technologies for recording the paths of moving objects have evolved (e.g., video, cell-phone, and GPS tracking) contemporary GIScientists have found new
opportunities for quantitative analysis using time geography with GISci (e.g., Miller 1991, Kwan 1998). Other quantitative methods for analyzing movement have stemmed from geography’s strong legacy in spatial point pattern analysis (e.g., Gao et al. 2010), as movement data are commonly represented by a sequence of points. Computational geometry has played a leading role in recent advances in analyzing movement data (e.g., Laube et al. 2005). As well, methods for representing movement data using areal data formats, for example polygons (Downs and Horner 2009) or fields (Downs 2010), remain ongoing research areas. The study of movement is of interest outside of GISci, for example wildlife ecology (Nathan et al. 2008), urban planning (Drewe 2005), and military applications (Wells 1981). Further, the study of movement has a long history in physics. Even Hägerstrand’s time geography was strongly influenced by the ideas of physicists from the early 20th century (Rose 1977, Hallin 1991). For example, the diagram of the space-time cone from time geography can be clearly related to the past and future light-cones used in Einstein’s relativity.

Movement is a complex process that operates through both space and time. Representing the temporal dimension in geographic studies has presented a challenge for GISci to move beyond static (map-based) representations of space (Chrisman 1998, Laube et al. 2007). Despite notable advances at incorporating temporal dynamics in GISci (e.g., Pultar et al. 2010), integrating the study of space and time remains at the forefront of GISci research, as evidenced by the special symposium on space-time integration in GISci at the 2011 annual meeting of the Association of American Geographers. How to effectively integrate time into the quantitative analysis of movement, specifically movement data stored in a GIS, is at the core of this review.
The growth of spatial methods for quantitative analysis of movement data has been facilitated by developments in movement databases that now provide efficient methods for storing, indexing, and querying movement data (Güting and Schneider 2005). Despite the large body of existing literature on the topic of moving object databases, it remains an active area of research as new tools (e.g., Güting et al. 2010a) and applications (e.g., Jensen et al. 2010) continue to develop. Data visualization methods have developed alongside these readily available movement databases; in GISci this practice is termed geovisualization (Dykes et al. 2005). Given the sheer volume of data often contained in movement databases, geovisualization can be a powerful tool for identifying patterns in movement databases – a process referred to as visual analytics (Thomas and Cook 2005). A complete treatment of either of these topics is beyond the scope of this review, and we restrict the contents of this review to, as the title suggests, those methods for analyzing movement data that are quantitative in nature. We would point those interested in more information on movement databases to the comprehensive book by Güting and Schneider (2005) and a recent special issue on data management for mobile services (VLDB Journal, 20(5), Güting and Mamoulis 2011). For those interested in more information on visual analytics for movement data we refer readers to Andrienko and Andrienko (2007), and to the special issue from IJGIS entitled geospatial visual analytics: focus on time (IJGIS, 24(10), Andrienko et al. 2010).

The objective of this review is to provide an unbiased evaluation of the usefulness and shortcomings of existing quantitative methods for movement data, while highlighting techniques that have particular merit with emerging movement datasets. Challenges to the development and application of quantitative methods with movement data are identified
in an attempt to locate avenues for future research. An outline of this article is as follows;

section 2 contains a brief introduction to the properties of movement data, and how

movement data is typically represented in a GIS. In section 3 we review the existing

literature on quantitative analysis of movement data separated into seven classes of

methods: 1) time geography, 2) path descriptors, 3) similarity indices, 4) pattern and

cluster methods, 5) individual-group dynamics, 6) spatial field methods, and 7) spatial

range methods. With section 4 we provide a discussion of the challenges routinely faced

in GISci when analyzing movement data and, what we feel are, some future directions for

quantitative movement analysis. Lastly, we close with some conclusions.

2 – Movement Data

Movement is a continuous process that operates in both the spatial and temporal

domains. Movement data are used to represent the continuous process of movement for

geographical analysis. Due to existing geospatial data collection and storage techniques,

movement data are most commonly represented as a collection of spatial point objects

with time stored as an attribute. A more formal definition of movement data is the

collection \( \{M_t\} \) of \( t = 1 \ldots n \) ordered records each comprised of the triple \( <ID, S, T> \),

where \( ID \) is a unique object identifier, \( S \) are spatial coordinates, and \( T \) a sequential (non-
duplicated) time-stamp (Hornsby and Egenhofer 2002). A number of terms are used

synonymously for movement data (see Table 1); here we use the term \( path \) to represent

the ordered sequence of records portraying individual/object movement, the term \( fix \)

when discussing a single record from a path, and the term \( movement database \) to describe

a collection of paths. The term \( movement data \) is used in broader contexts when
discussing the study of movement, to refer generally to fixes, paths, and movement databases.

While movement data have historically been collected using a variety of techniques, most current acquisition schemes use some form of wireless sensor (e.g., GPS, cellular phone records, radio telemetry). Calenge et al. (2009) identify two types of sampling commonly employed in the collection of movement data – regular and irregular. Regular paths are those where fixes are acquired at an even temporal interval, for example recording one fix per minute. Irregular paths are those where fixes are acquired at unequal temporal intervals, for example paths collected from cell phone call records. The term granularity is used to refer to the resolution of a path (Hornsby and Egenhofer 2002). Finer granularities are associated with frequent sampling intervals, and provide a detailed representation of movement. Conversely, coarser granularities correspond to sparse sampling and less-detailed representation of movement.

Technological developments now facilitate finer sampling intervals in movement paths (e.g., 1 fix / second), and movement data can be used to represent a (near) continuous movement path (Laube et al. 2007). However, these sensor-specific sampling designs may not be suitable for all analysis questions, requiring the use of re-sampling (up- or down-sampling) to fit a given research need (see Turchin 1998, and Hornsby and Egenhofer 2002 for a more thorough discussion of changing granularity).

Spaccapietra et al. (2008) present an alternative view of movement data granularity, defining a path as consisting of stops and moves separating a path into periods of movement and stationary behavior. This conforms with the event-based model
for movement data outlined by Stewart Hornsby and Cole (2007) which contrasts with
the coordinate-based representation of movement typically employed. An event based
model for movement data still allows for the detection of movement patterns, but with
focus placed on combinations or sequences of events that identify a specific behavior,
such as an exodus of objects out of a zone or region (Stewart Hornsby and Cole 2007).
Further, event based models allow for enriching movement data with the geographic
information associated with events, for instance if events are related to spatial regions the
attributes of each region.

3 – Review of Methods

This section contains a review of quantitative analysis methods that exist within
seven areas of movement research; 1) time geography, 2) path descriptors, 3) path
similarity indices, 4) pattern and cluster methods, 5) individual-group dynamics, 6)
spatial field methods, and 7) spatial range methods. We emphasize techniques we feel
have particular merit for analysis with novel and emerging movement datasets.

3.1 – Time Geography

The concept of time geography was first presented in the 1960’s and 1970’s by
Torsten Hägerstrand at the Research Group for Process and System Analysis in Human
Geography at the University of Lund, Sweden (Lenntorp 1999). Time geography
(Hägerstrand 1970) represents a framework for investigating the constraints, such as an
object’s maximum travel speed, on movement in both the spatial and temporal
dimensions. Hägerstrand expanded on the purely physical limitations of movement,
identifying three other types of constraints: capability, coupling, and authority
Constraints. Capability constraints limit the activities of the individual because of their biological construction and abilities, for example the necessity to eat and sleep. Coupling constraints represent specific locations in space-time an individual must visit that limit movement possibilities. Authority constraints are opposite of coupling constraints, locations in space time an individual cannot visit, for example a mall after it has closed. Contemporaries expanded on Hägerstand’s work providing both theoretical (Parkes and Thrift 1975, Pred 1981) and applied (Lenntorp 1976, Burns 1979) extensions. Originally, time geography was used solely to investigate the movement of humans, but has since been reformulated for use with transportation networks (Miller 1991) and wildlife ecology (Baer and Butler 2000).

Time geography uses volumes (Figure 1) capable of capturing the movement limits of an object. A 3-D space (often termed cube, Kraak 2003, or aquarium, Kwan 2004), with two spatial axes representing geographic space and a third orthogonal axis for time, is used to develop time geography volumes. The space-time cone (Figure 1a) identifies the future movement possibilities of an object. A space-time prism (Figure 1b) is used to quantify movement possibilities between known start and end locations. The potential path area is the projection of the space-time prism onto geographic space (Figure 1c), and is a purely spatial measurement of movement capability. A path is used to portray the trajectory of movement through space-time. Bundling (Figure 1d) occurs when multiple paths coincide in space and time, for example taking the same bus to work. Typically, time geography is discussed qualitatively in terms of the aforementioned volumes, but Miller (2005) has provided mathematical definitions for time geography concepts that can be used in more rigorous quantitative analyses.
Recently, with advances in GIS and movement data, time geography is experiencing a resurgence (Miller 2003). Lenntorp (1999) explains how time geography has reached ‘the end of it’s beginning’, suggesting that current and future research using GIS and novel movement datasets will present new and exciting developments in time geography. Examples include using time geography to investigate mobility data on a network (Miller and Wu 2000), factoring in uncertainty (Neutens et al. 2007), field-based time geography (Miller and Bridwell 2009, further discussed in S3.6), and the development of a probabilistic time geography (Winter 2009, further discussed in S3.6).

Time geography represents a useful tool for quantitative analysis of movement as it contains a framework for measuring space-time bounds on movement. Movement models that fail to consider the constraints provided by space and time often result in misleading conclusions (Long and Nelson 2012). Methods that explicitly consider time geography principles, even unknowingly (e.g., Yu and Kim 2006), avoid such deceptions.

3.2 – Path Descriptors

Path descriptors are measurements of path characteristics, for example velocity, acceleration, and turning azimuth. Typically path descriptors may be calculated at each point in a movement dataset, and can be scaled appropriately to represent interval or global averages. Dodge et al. (2008) categorize a number of path descriptors as primitive parameters, primary derivatives, or secondary derivatives based on simple measurements in space, time, and space-time (see Table 2). Ecologists routinely use simple path descriptors in the study of wildlife movement (Turchin 1998). Measures of movement tortuosity have also been developed for the study of wildlife, for example path entropy
(Claussen et al. 1997), sinuosity (Benhamou 2004), and fractal dimension (Dicke and Burrough 1988). Related to these are stochastic movement models (i.e., models where fixes are obtained via random draws from distributions for movement displacement and turning angle) such as Lévy flights (Viswanathan et al. 1996) and correlated random walks (Kareiva and Shigesada 1983). When movement data are statistically fit to such models, interpretation of model parameters can provide useful quantitative inference.

<approximate location Table 2>

3.3 – Path Similarity Indices

Path similarity indices are routinely used to quantify the level of similarity between two movement trajectories. It is desirable for similarity indices to take the form of a metric distance function, as metric functions are able to distinguish objects on an interval scale of measurement (Sinha and Mark 2005). A metric distance function \( d \) is one that computes a generalized scalar distance between two objects while satisfying the following four properties (Duda et al. 2001):

(i) Non-negativity: \( d(x, y) \geq 0 \);
(ii) Reflexivity (uniqueness): \( d(x, y) = 0 \), iff \( x = y \);
(iii) Symmetry: \( d(x, y) = d(y, x) \);
(iv) Triangle Inequality: \( d(x, z) \leq d(x, y) + d(y, z) \)

The simplest similarity metric is a Euclidean measurement. Sinha & Mark (2005) implement a time-weighted distance metric where spatial proximity (Euclidean) is weighted by its temporal duration. Sinha & Mark (2005) also present a modified version of the time-weighted distance metric for the situation where the two objects move over different time intervals. Because the time-weighting is based on the duration an object spends at a given spatial location, this index works best with movement data defined as a series of stops and moves such as suggested by Spaccapietra et al. (2008). Yanagisawa et
al. (2003) present an alternative Euclidean-based similarity index that focuses on the shape of the movement path by normalizing the spatial coordinates of a path to a common plane. Euclidean measurements in the normalized spatial plane are used to identify similarly shaped movement paths. Euclidean distance is appropriate for comparisons in the spatial or temporal domains. However, Euclidean measurements are limited when data are represented with different scales (spatial and temporal). That is, what is the temporal equivalent to a 1 km distance in space? Despite these limitations, Euclidean distance similarity indices are frequently implemented by fixing either space or time and considering Euclidean distance in the other dimension, such as the above examples.

Other distance metrics may be more appropriate for assessing path similarities. The Hausdorff distance is a shape comparison metric commonly used to evaluate the similarity of two point sets (Huttenlocher et al. 1993), which has also been used to measure the similarity of movement paths. Given two movement paths $M^a$ and $M^b$, the Hausdorff distance is defined as:

\[
H(M^a, M^b) = \max \left( h(M^a, M^b), h(M^b, M^a) \right) \qquad [1]
\]

with

\[
h(M^a, M^b) = \max_{t \in T} \left( \min_{s \in S} d(M^a_t, M^b_s) \right) \qquad [2]
\]

where $t$ and $s$ are used to index fixes from $M^a$ and $M^b$ respectively, and $d$ is a distance operator (e.g., Euclidean). Not originally designed for movement data, the Hausdorff distance performs poorly when analyzing movement paths as it fails to consider the ordering of points (Zhang et al. 2006), and is sensitive to outliers and data noise (Shao et al. 2010). As such, modified versions of the Hausdorff distance metric have been
designed specifically for use with movement paths (e.g., Atev et al. 2006, Shao et al. 2010).

The Fréchet distance metric may be more appropriate as a path similarity index as it was initially designed for comparing polygonal curves. Formally the Fréchet distance for two movement paths $M^a$ and $M^b$ is defined as:

$$
\delta_f (M^a, M^b) = \inf_{\alpha, \beta} \max_{t, s \in [0,1]} d(M^a(\alpha(t)), M^b(\beta(s)))
$$

Where $\alpha$ (resp. $\beta$) is an arbitrary continuous non-decreasing function from $[0,1]$ onto $[t_1...t_n]$ (resp. $[s_1...s_n]$) and $d$ is a distance operator (Alt and Godau 1995). In simple terms, the Fréchet distance measures the maximum distance apart of two coinciding movement paths. The Fréchet distance, is best conceptualized using the analogy of a person walking their dog, where no backwards movement is allowed. In the dog walking example, the Fréchet distance is the minimum length of the dog’s leash. The discretized form of the Fréchet distance metric (Eiter and Mannila 1994) is useful for its computation with movement data collected by discrete fixes, as described in section 2. In applications involving objects that move with the same temporal granularity this calculation is simply the maximum distance in space between any pair of fixes taken at the same time. However, when object movement is recorded at differing temporal granularities or extents, the value of the Fréchet distance metric is through the use of the scaling functions ($\alpha$, $\beta$) to measure similarity.

Vlachos et al. (2002) use longest common subsequences (LCSS), a method taken from time-series analysis, to identify similar movement paths. The LCSS is defined as the number of consecutive fixes from two (or more) paths ($M^a, M^b, ...$) that are within $d$ spatial and $\tau$ temporal units of each other. This method can be extended to paths that
move at a distance, using mapping function $f(M)$ to translate $M^b$ onto a space equivalent
to $M^a$. LCSS is advantageous as it is able to address issues relating movement paths taken
at different temporal granularities and/or extents. LCSS is efficient even with paths that
contain a significant amount of data noise. When outlying fixes are likely to influence the
calculation of other similarity indices LCSS is advantageous as it is insensitive to
extreme outliers. The disadvantage of the LCSS method is that it relies on the subjective
definition of thresholds – $d$ and $\tau$, and it fails the triangle inequality test (iv. above), and is
therefore not a metric distance function.

Similarity indices have also been extended to objects moving along a network.

For example, Hwang et al. (2005) calculate similarity using points-of-interest, such as
major intersections. Movement paths are considered similar if they pass through the same
points-of-interest in the same order. This index is not a metric distance function, but
moves away from Euclidean based measurements which are inappropriate in a network
scenario.

Recently, a new similarity method has been proposed by Dodge et al. (2012).
Here, a movement path is separated into segments where specific movement parameter
patterns (and derivatives of) are observed. In their example, velocity is the parameter of
interest, and the metrics deviation from the mean and sinuosity are used to define
movement parameter classes. For example, the letters A-D could be used to denote 4
unique movement parameter classes, and a path could then be represented as the
sequence [ACBCACBDBDA]. To assess the similarity of two paths, a modified version
of the edit distance (a string matching algorithm) is computed on the movement
parameter class sequences. This method measures similarity in the selected movement
parameters, rather than in the space-time geometry of the movement paths. As such, it may be more appropriate when similarity in various parameters, rather than space-time geometry is specifically of interest, for instance, in the study of hurricane path dynamics, as demonstrated by Dodge et al. (2012).

When objects interactively move with each other at a distance, they often exhibit correlated movement. Typically, similarity indices may identify such correlated movements by mapping the spatial coordinates of one path onto the spatial plane equivalent to the other. Alternatively, Shirabe (2006) presents a method for computing the correlation coefficient between two movement paths, each represented as a vector time-series. Consider a path $M$ with $t = 1 \ldots n$ fixes, then for $t = 2 \ldots n$, $V = [M_t - M_{t-1}] = [v_t]$, is a vector time series of $M$. Given two movement paths ($M^v$, $M^w$) represented as vector time-series $V$ and $W$, the correlation coefficient is defined as:

$$r(V, W) = \frac{\sum_{t=1}^{n-1} (v_t - \bar{v}) \cdot (w_t - \bar{w})}{\sqrt{\sum_{t=1}^{n-1} |v_t - \bar{v}|^2} \sqrt{\sum_{t=1}^{n-1} |w_t - \bar{w}|^2}} \quad [4]$$

Where $\bar{v} = \frac{1}{n-1} \sum_{t=1}^{n-1} v_t$ (resp. $\bar{w}$) are mean coordinate vectors of $(V, W)$. Note that a movement path of $n$ fixes is comprised of $n-1$ movement vectors, this distinction we keep for consistency with other methods. The numerator in [4] is the covariance, which indicates how the two motions deviate together from their respective means (Shirabe 2006). Geometrically, the dot product in the numerator is the product of vector lengths multiplied by the cosine of the angle between them, which can be interpreted as the similarity. The correlation index ranges from -1 to 1, identifying both negatively and positively correlated movements. Important to note is that this correlation coefficient
relies on each movement’s deviation from the respective mean, not the raw values of each observed movement. Relating correlations to a global mean can be advantageous in cases where two movements are correlated, but do not move in the same direction. The first drawback of the formulation in [4] is that we are unable to disentangle the effects of correlation in azimuth vs. magnitude of movements. A metric decomposed into each of these components would be advantageous in situations where such distinctions are necessary. A second drawback of equation [4] is that it requires that the fixes from each movement path be taken simultaneously in order to be valid, which is not always realistic. However, Shirabe (2006) does present an extension for modifying [4] to measure movement path correlations at a temporal lag.

3.4 – Pattern and Cluster Methods

Many applications are interested in identifying broad spatial-temporal patterns from large movement databases (Benkert et al. 2007, Palma et al. 2008, Verhein and Chawla 2008). For example, in the study of tourist behavior, often the goal is to identify places of interest that are frequently visited (e.g., Ahas et al. 2007). Alternatively, studying commuter patterns typically involves the identification of intersections and routes being used by multiple individuals (Verhein and Chawla 2006). In these situations, pattern and cluster methods are employed to identify similar movement behaviors or places of interest.

Early work on indexing and querying movement databases coming from the computer and database science literature (e.g., Güting et al. 2000, Pfoser et al. 2000) has been essential to the development of pattern and cluster methods. For instance, many methods for identifying patterns and clusters in large movement databases implement a
simple spatial or temporal query (Erwig et al. 1999). Alternatively, pattern or cluster
methods may implement one of the aforementioned path similarity indices and perform
pair-wise similarity computations over all permutations of stored movement paths. Paths
identified as similar based on a query or similarity index may convey some movement
pattern, or belong to the same cluster. The use of the term ‘cluster’ comes from methods
for statistical analysis of spatial point patterns (Diggle 2003), as many approaches used in
point pattern analysis have been adopted for movement data. For example, both Gao et al.
(2010) and Güting et al. (2010b) describe methods for performing $k$-nearest neighbor
queries in movement databases.

For the most part, the identification of patterns and clusters in large movement
databases focus on one of space, time, or space-time. Methods that identify spatial
clusters look at space first and time second, if at all (e.g., Benkert et al. 2007). The
simplest methods for detecting spatial clusters in movement databases generally require
that fixes from individual paths be represented as spatial points. Other spatial methods
look to define regions of interest (static or dynamic) and identify times at which
movement fixes are clustered in these spaces (Giannotti et al. 2007). Alternatively,
temporal clusters look at time first and space second, (e.g., D'Auria et al. 2005, Nanni and
Pedreschi 2006). Temporal clustering is enhanced (Palma et al. 2008) when movement
paths are represented by a sequence of stops (representing activities) and moves
(Spacapietra et al. 2008).

Space-time approaches to identifying patterns and clusters strive to consider space
and time simultaneously. This is difficult, as previously mentioned, due to scaling
differences between space and time. Most space-time approaches fail to properly scale
space and time and degenerate to spatial clustering methods linked through time (e.g., Kalnis et al. 2005). Such methods routinely consider the following problem: given \( p \) mobile objects, \( M^i, i = 1 \ldots p \). Each \( M^i \) consists of \( n \) fixes taken at coinciding times \( t = (1, \ldots n) \). A set of \( \alpha (1 \leq \alpha \leq p) \) spatial clusters are identified at each time \( t \) (for example with multivariate clustering) using the spatial \((x, y)\) coordinates of \( M^i(t) \). In one example, Shoshany et al. (2007) link clusters through time using linear programming. In their example, moving objects \( M^i \) can switch between clusters, but all \( M^i \) must belong to a cluster, as well clusters can emerge or disappear over time. The appeal of this approach is that linear programming, frequently used in optimization research, can identify flows and trends in movement data clusters.

Spatial-temporal association rules (STAR) learning represents an algorithm-based method for discovering spatial-temporal patterns in movement databases (Verhein and Chawla 2006, 2008). The patterns found by STAR methods are able to identify sources, sinks, and thoroughfares in large mobility databases. Verhein and Chawla (2008) demonstrate a STAR-miner software that implements their algorithm, and apply it to a caribou dataset. STAR patterns rely on pre-determined spatial units (termed regions) over which the algorithm is run. Unfortunately, the use of explicit spatial regions in their derivation means that STAR are especially sensitive to changes in the definition of regions (known as the modifiable areal unit problem - Openshaw 1984).

Pattern and cluster methods for movement data have also drawn on existing methods from other applications. Shoval and Isaacson (2007) propose sequence alignment methods, originally used to analyze DNA, as a way to identify patterns in human travel behavior. With movement data, sequence alignment methods are able to
identify groups of objects that follow a similar sequence of events (e.g., using an event
based movement data representation, as in Stewart Hornsby and Cole 2007). Shoval and
Isaacson (2007) apply sequence alignment methods to tourist movement data and
conclude that sequence alignment methods have potential for identifying patterns of
spatial behavior in large movement databases. In another example, Eagle and Pentland
Eigenbehaviors represent trends or routines in individual movement data. Principle
component analysis is used to identify the eigenbehaviors of each person in their dataset.
In their example using the movements of people’s daily routines, three trends emerge:
workday, weekend, and other behaviors. Increasingly complex questions could be
addressed using the eigenbehavior method.

3.5 – Individual-Group Dynamics

The term individual-group dynamics is used to classify a suite of methods that
focus on individual object movement within the context of a larger group. This differs
fundamentally from methods for identifying patterns and clusters in movement databases.
Most current methods for investigating individual-group dynamics rely on computational
algorithms capable of searching movement databases for specific, pre-defined patterns.
These algorithms are often computationally demanding and inefficient (Gudmundsson et
al. 2007), and thus primarily used only in small, case-study examples.

Laube et al. (2004, 2005) provide the most comprehensive examination of
individual group-dynamics. Their concept of relative motion (REMO) can be used to
detect specific patterns (constancy, concurrence, and trend-setters) in groups of moving
objects. Constancy represents when an object moves in the same direction for a number
of consecutive fixes. An episode of concurrence occurs when multiple moving objects move in the same direction at the same time. Trend-setters are objects that move in a given direction ahead of a concurrence episode by a group of objects. Trend-setting is identified as the most interesting property, and examined in more detail using the sport of soccer as an example. Players that exhibit trend-setting behavior are able to better anticipate the movement of play. Their concept of trend-setting has been further developed for identifying leaders and followers in groups of moving objects, which is potentially useful for the analysis of wildlife movement data (Andersson et al. 2008).

Laube et al. (2005)’s REMO method uses only movement azimuths to determine relative motion. All other movement attributes, such as speed or distance, are ignored in their derivation. Thus, REMO is useful only in situations where a group of objects move with similar speeds and are contained in a relatable geographic space, such as the soccer example. Another disadvantage is that the REMO method relies on the definition of azimuthal breakpoints to define when objects are moving in a similar direction (e.g., East is between 45° and 135°). Due to their discreteness, these breakpoints can lead to misleading interpretations, for example when objects move in similar directions on either side of a breakpoint. Alternatively, Noyon et al. (2007) evaluate the relative movement of objects from the point-of-view of an observer within the system. Using changes in relative inter-object distance and velocity, Noyon et al. (2007) identify relative behavior, for example collision avoidance. Furthermore, Noyon et al. (2007) suggest that such relative movement behavior also include other regions-of-interest such as lines and polygons, which they include in their derivation.
Another problem routinely encountered in the study of movement is the detection of flocks and convoys (e.g., groups of individuals that move as a cohesive unit). A flock (see Figure 2a) is defined as a group of at least \( m \) moving objects \((M)\) contained within a circle of radius \( r \) over a minimum time interval - \( \tau \) (Gudmundsson and van Kreveld 2006, Benkert et al. 2008). Alternatively, a convoy (see Figure 2b) is defined as a group of at least \( m \) moving objects \((M)\) that are density connected at a distance \( d \) over a minimum time interval - \( \tau \) (Jeung et al. 2008). Density connected implies that there exists a sequence of segments connecting all points in the convoy, each segment with length \( \leq d \).

This definition of convoy relaxes the circular requirement of flocks affording flexibility in the shape and extent of convoys that can be identified, for example Canada geese forming their characteristic V-shape. Methods that look at flock/convoy behavior have obvious usefulness in the study of wildlife herds, but also in monitoring crowd dynamics at large events (Benkert et al. 2008). Like space-time clustering, methods describing flocks or convoys build upon Hägerstrand’s concept of bundling, identifying areas where objects move coincidentally in space-time. The fundamental difference between the identification of flocks or convoys and space-time cluster methods is that the definition of a flock or convoy explicitly considers the individual in relation to the group in its definition. That is, focus is placed on membership to a given group, with explicit consideration of minimum requirements for flock or convoy behavior (e.g., the parameters \( m \) and \( \tau \)). Space-time cluster methods focus more on identifying broader patterns, typically from large movement databases, and generally rely on pair-wise comparisons of individual movement paths.

<approximate location Figure 2>
Recently, free space diagrams have been proposed for identifying single-file motion in movement databases (Buchin et al. 2010). To conceptualize a free space diagram consider two movement paths \((M^a, M^b)\), over the time intervals \(m\) and \(n\) respectively, where the trajectory between fixes is given by some linear or other model (e.g., Tremblay et al. 2006). The functions \(\varphi_a\) and \(\varphi_b\) give the position of the objects \(a\) and \(b\) at time \(t\). The free space diagram for \(a\) and \(b\) (following Buchin et al. 2010) is given by:

\[
 F_\delta(M^a, M^b) = \{ (t^a, t^b) \in [1, n] \times [1, m] : \|\varphi_a(t^a) - \varphi_b(t^b)\| \leq \delta \} \quad [5]
\]

which defines the set of all points in \(\varphi_a\) and \(\varphi_b\) that have a Euclidean distance below some threshold \(\delta\). The map of \(F_\delta\) describes a two dimensional space where the axes correspond to the two paths, and the free space is defined as anywhere along the paths where the distance between the two paths is below the threshold \(\delta\). Buchin et al. (2010) demonstrate a method for interpreting free-space diagrams capable of identifying single-file movement patterns in groups of moving objects. A criticism of this method is that it relies on a subjectively defined threshold \(\delta\), to constrain the single-file movement process. Single-file motion has intuitive meaning, but is especially difficult to conceptualize geometrically. Methods that use Euclidean geometry to measure the spatial separation between leaders and followers (e.g., Andersson et al. 2008) are inadequate for identifying single-file movement warranting the free-space diagram approach.

### 3.6 – Spatial Field Methods

Often it is of interest to represent a movement path (or many movement paths) as a spatial field in order to identify areas in space (or space-time) that are more or less frequently visited. Field based representations are especially useful for visualizing large quantities of movement data when maps become cluttered. As many other spatial datasets...
are stored as raster fields, a field-based representation of movement allows quantitative
map comparisons to be performed in a GIS.

Most methods for representing movement data as spatial fields have evolved from
those used to analyze spatial point patterns. When spatial point pattern methods are
employed the temporal component of movement fixes is ignored. Spatial point pattern
methods can be separated into quadrat or density based methods (Diggle 2003). The
simplest quadrat methods involve subdividing a study area into a regular grid and
determining point densities within each cell (e.g., Dykes and Mountain 2003,
Hadjieleftheriou et al. 2003). Cells with high point densities indicate spatial locations of
high use. Hengl (2008) proposes a quadrat based space-time density measure based on
distance and velocity within each cell [6].

\[
D_{xyt}(j) = \frac{\hat{d}_j}{\hat{v}_j} \quad [6]
\]

Here \(D_{xyt}(j)\) is the space-time density at cell \(j\), \(\hat{d}_j\) is the length of the movement path
within cell \(j\), and \(\hat{v}_j\) is the average velocity of movement within cell \(j\). For a single
moving object the space-time density is simply interpreted as the duration of time the
object spends within each cell. If calculated for a movement database of many objects,
areas with higher space-time densities represent those where more objects spend more
time, the opposite with low values (Hengl et al. 2008). This approach has been extended
for three-dimensional visualization, where density is related to the lengths of multiple
paths in 3-D voxels defined by two spatial dimensions and a temporal dimension
(Demšar and Virrantaus 2010). Voxel densities are visualized in a space-time cube
(aquarium), and can be used for exploratory analysis of large movement databases.
Density based methods in spatial point pattern analysis stem from bivariate probability models, where movement fixes represent sampled locations from a two-dimensional probability density function (Silverman 1986). In the analysis of wildlife, density based models are frequently used to generate estimates of animal space use (also discussed in S3.7). Worton (1989) first applied kernel density estimation (KDE) to wildlife movement data to derive such a surface, termed a *utilization distribution* (Jennrich and Turner 1969). In movement applications, KDE can be interpreted as the intensity of space use based upon a collection of fixes. Calculation of KDE requires selection of a kernel shape and bandwidth parameter, with no consensus on the best way to do so (Hemson et al. 2005, Kie et al. 2010). Alternatively, Downs (2010) has proposed time geography’s potential path area (see Figure 1) to replace the kernel shape and bandwidth parameter, representing a novel approach for integrating temporal constraints into KDE analysis. Downs (2010) replaces the traditional kernel function with one based on the potential path area (termed geo-ellipse – $G$) from time geography [7].

$$f_j(x) = \frac{1}{(n-1)(t_j - t_i)v^2} \sum_{i=1}^{n-1} G \left( \frac{\|x - M_i\| + \|M_j - x\|}{(t_j - t_i)v} \right) [7]$$

The numerator in this function sums the distance between a given point $x$ and the object’s locations ($M$) at times $i$ and $j$. The denominator is the maximum distance the object could have travelled in that time interval given its maximum velocity – $v$. Others have seen the need to move away from continuous representations of space, and have developed KDE for networks (Borruso 2008, Okabe et al. 2009). Such analysis is more appropriate for depicting the movement of urban travelers as their movement is restricted to travel networks of roads, paths, and sidewalks.
Random walks and diffusion theory have also been used to model movement as a continuous spatial field. Horne et al. (2007) use Brownian bridges to model wildlife movement as a continuous probability surface. Between two consecutive mobility points, the probability an object is at a given location at time $t$ is defined using a bivariate normal probability density function. More recently, probabilistic time geography has been proposed (Winter 2009), where a similar probability surface is based on discrete random walks in a cellular automata environment. Winter & Yin (2010) extend on the ideas of Winter (2009) to include directed movements. Random walks are used to derive a probability surface which explicitly considers the time geographic constraints on object movement, using a similarly defined bivariate normal probability surface. Both Winter & Yin (2010) and Horne et al. (2007) discuss the fact that determining movement probabilities based on random walks is limited when objects do not move randomly.

Future work looking at probabilistic movement using other movement models (e.g., correlated random walks or on a network) is thus warranted for moving objects that can be modeled this way. Alternatively, Miller & Bridwell (2009) propose a field-based time geography. Field-based time geography uses movement cost surfaces in the calculation of time geography volumes. Movement possibilities are evaluated in a similar manner to Winter and Yin (2010) but based on an underlying movement cost surface (e.g., as in least-cost path analysis in GIS, Douglas 1994). This approach is advantageous in that it directly considers underlying variables impacting movement, however is limited in that an accurate cost surface must be derived.

Brillinger et al. (2001, 2004) provide a unique approach for discovering patterns in movement data. Stochastic differential equations are used to model movement as a
Markov process. The drift term in the stochastic movement model can be interpreted as a spatial velocity field and used for exploratory analysis. The spatial velocity field represents a potential function, whereby points of attraction and repulsion can be identified. Methods for statistical inference (e.g., jackknifing) can be used to identify statistically significant movement patterns within this velocity field (Brillinger et al. 2002). Brillinger (2007) further applies this approach for analyzing the flow of play in soccer, where the spatial velocity field for ball movement is used to investigate a team’s attack formation.

3.7–Spatial Range Methods

Spatial range can be broadly defined as the area (generally represented as a polygon) containing an object’s movement. Measures of spatial range can be useful for examining object mobility and space use. Aspatial metrics, such as net displacement (Turchin 1998), provide no information on the spatial distribution of movement, simply measuring distance, thus spatial measurements are warranted. Furthermore, researchers are often interested in intersections and/or differences in movement ranges (e.g., Righton and Mills 2006). In such cases it is advantageous to represent point/line movement data in an areal format (e.g., as a polygon).

The practice of representing movement data using spatial polygons has been developed primarily by wildlife ecologists for studying wildlife home ranges (Burt 1943), however, the concept of home range has also been applied to other subjects (e.g., children, Andrews 1973). Spatial range methods typically rely on the geometric properties of movement data, for example the calculation of the minimum convex polygon, a common measure of wildlife home range (Laver and Kelly 2008). Other
geometric methods include harmonic mean (Dixon and Chapman 1980), Voronoi polygons (Casaer et al. 1999), and characteristic hull (Downs and Horner 2009). It is also common to extract spatial range polygons from spatial field representations of movement (e.g., those from S3.6) by extracting polygon contours based on density. For example, with KDE a 95% volume contour is frequently used to delineate wildlife home range, while a 50% volume contour is used to delineate core habitat areas (Laver and Kelly 2008). These spatial range methods ignore temporal information stored in movement data and are likely to contain areas never visited by the object (commission error), and miss actually visited locations (omission error) (Sanderson 1966).

Time geography volumes may also be used for generating spatial range estimates. Long & Nelson (2012) propose a spatial range method for wildlife movement data based on time geography’s potential path area (Figure 1c). This method is capable of identifying omission and commission errors in other spatial range methods (Long and Nelson 2012). Such time geographic analysis is commonly used to study accessibility in the context of human movement (Kwan 1998). The value of the potential path area as a spatial range method is that it explicitly considers the temporal sequencing of movement data in a time geography context. Spatial range methods that consider the temporal component of movement data are advantageous over purely spatial methods (such as convex polygons) as they consider movement data as a sequence of spatial points taken through time, rather than as an arbitrary collection of spatial points.

4 – Discussion

4.1 – Time
The first and foremost challenge to the quantitative analysis of movement data is how to effectively characterize time. Despite having well-developed theory and tools for analyzing space, geographers and the GIS community have historically struggled with the temporal dimension (Peuquet 1994). Time is a single, continuous dimension that can be portrayed as either monotonically linear or cyclical (Frank 1998). If time is portrayed as linear, objects are not capable of re-visiting instances in time. If time is portrayed as cyclical, the beginning of a new cycle infers that time is reset to some initial state, thus revisiting is facilitated. For example, consider research on human daily routines; within each day time is treated linearly, but is reset at the beginning of each day signifying the start of a new cycle. Movement data collected over long periods may contain both linear and cyclical temporal patterns, confounding representation and analysis.

Theoretical constructs for including time in GIS have long been discussed (Langran and Chrisman 1988, Peuquet 1994) but remain challenging. Some spatial datasets are easily represented at discrete time intervals in a GIS as different layers, for example land cover data in different years. This representation allows for vertical analysis through time using relatively simple map algebra (Mennis et al. 2005). Vertical analysis through time is not straightforward with movement data, as objects move in both space and time and cannot be explicitly linked through the spatial dimension. Others have suggested the notion that geography’s fetish for the static (Raper 2002) may lie at the root of the time problem. In practice, researchers have begun to use a 3-D aquarium (drawing on Hägerstrand’s ideas) for representing time in GIS, however this is principally a visualization tool (e.g., Kraak 2003, Andrienko and Andrienko 2007, Shaw et al. 2008). Dynamic views (i.e., animations) may overcome the drawbacks of static portrayals of
movement, allowing more fluid representations of velocity and acceleration properties
(Andrienko et al. 2005). However, dynamic views are also visual-based, and lack
potential for developing quantitative analyses.

The challenge has been finding appropriate ways to simultaneously represent the
different scales of measurement for temporal and spatial attributes associated with
movement. Consider that it is common to use measurements of time and space
interchangeably in queries associated with movement from everyday life, for example if
you were asked the question: how far is it from here to the grocery store? You might
answer with “about 2 kilometers” or alternatively with “about a 5 minute drive”. Here, a
question of spatial distance associated with movement can be equivalently answered
using a spatial measurement (2 km) or temporal measurement (5 minutes). This has led to
alternative conceptualizations of movement where space and time can be represented
using relationships that can scale from spatial to temporal measurements, and vice-versa
(Parkes and Thrift 1975). For example, travel can be considered as the consumption of
physical distance through time (Forer 1998). However in the previous scenario, you may
have also answered with “about a 5 minute drive, depending on traffic”. Alternatively,
one might add that it depends on mode of transport (e.g., whether you walk or drive).
This alternative view demonstrates the non-linear and dynamic relationship that exists
between space and time which confounds the direct exchange of measurements of space
and time (Forer 1998). With movement data, time is often stored alongside spatial
attributes (e.g., \( <x, y, t> \)), which naturally lends itself to Euclidean-type measurements in
the space-time aquarium. However, as demonstrated, time is poorly represented by such
direct physical measurements, because time cannot be represented as a linear function of
space. As there is still no consensus on the best way to represent time with movement
data, research on how to effectively characterize space and time in movement data
continues to require development.

Distance in space is easily computed using Euclidean (or other, such as network)
measurements. Differences in time are generally measured using clock times. The
conceptualization of a single space-time proximity measure remains one of the biggest
hurdles with quantitative analysis of movement data. Moving forward it is imperative to
go beyond simple Euclidean based measures, as time and space do not operate on equal
scales (Peuquet 2002). The Fréchet distance (Alt and Godau 1995) is an example of a
novel method for comparing the similarity of two movement paths that may prove useful
in future analyses. Nearest neighbor computations (e.g., Gao et al. 2010), most useful
with movement data stored as points, may also provide avenues for exploration.
Normalizing different data scales, common to other branches of quantitative analysis
such as multivariate cluster analysis (Duda et al. 2001), may be useful for comparing
movement processes across scales and relates to work using fractals for describing
movement datasets (Dicke and Burrough 1988). Normalization, however, may mask
scale specific patterns, and should be done with caution only when scale specific
behavior is less-important. Fundamentally, space and time have different dimensions and
require special consideration when analyzed together.

4.2 – Scale

With any spatial analysis the selection of analysis level (scale) will influence the
outcome of quantitative measures and the resulting inferences and conclusions (Dungan
et al. 2002). The study of scale and its impacts in spatial analysis remains a key topic in geographic studies. In the analysis of movement data Laube et al. (2007) identify four levels of analysis: instantaneous, interval, episodal, and global (Figure 3). The instantaneous (“local”) level represents measures computed at any point along a movement path. Interval (“focal”) level analysis takes the form of a moving temporal window, but may also use a moving spatial window. Episodal (“zonal”) level analysis looks at specific partitions of movement data, often related to some known event. Most common is global level analysis, where a movement dataset is represented as a complete path, from beginning to end, as a single entity. While some methods are specifically designed for a given level of analysis others can be applied to various levels. Methods that can be applied at different analysis levels may not scale from one level to the next, meaning results at a lower level may not sum to the global result, as is the case with some spatially local statistics (termed LISA - Anselin 1995).

Quantitative methods are also sensitive to changes in the temporal granularity at which movement data is represented (Laube and Purves 2011). Methods for changing granularity can be used when process scale is explicitly known, however this is rarely the case. When movement data are over-sampled (i.e., too fine a granularity) data noise can mask broader-scale process signals. When movement data are under-sampled (i.e., too coarse a granularity) important movement events are missed, leading to incorrect parameter estimates. Some ecologists have suggested that movement data should not be sampled at even time intervals, but rather as a sequence of moves or steps relating to individual behavior (Wiens et al. 1993, Turchin 1998). This aligns with the view of
Spaccapietra et al. (2008) that human movement data are best represented as a series of stops (representing activities, as in the event-based model of Stewart Hornsby and Cole 2007) and moves. However, many developed methods tend to perform better when implemented with regularly sampled movement data (e.g., Downs et al. 2012). As the toolbox of methods for the quantitative analysis of movement grows, it will be important to identify at what analysis level(s) and over which temporal granularities various methods are appropriate.

As previously identified, and following from Laube et al. (2007) and Laube and Purves (2011), there are two fundamental issues of scale associated with movement analysis, that is, analysis level and temporal granularity. Laube and Purves (2011) suggest a third issue of scale may also exist, in that many approaches for movement analysis are tested only on small, idealized datasets, and do not perform as expected when carried out on larger, real-life datasets. As a result, many existing methods cannot be readily implemented in practical scenarios with large volumes of movement data. We take an alternative view on this issue. Testing of methods with smaller, idealized datasets limits the scope of movement analysis to realistic and manageable problem sets, which are in turn appropriate with subsets of a larger movement database. For example, the detection of trend-setters (Laube et al. 2005) is only useful if there is some expectation about where, if observed, this pattern is meaningful. In applied research, one should be able to identify specific scenarios, within a larger movement database, where a given technique is appropriate. Once these specific scenarios are identified, for example using spatial-temporal queries, apply the technique of interest on this subset of the movement database. The result is a multi-tiered analysis, where a specified technique is only
performed on smaller, appropriate subsets of the data. The goal being to break down
larger movement datasets into pieces resembling the idealized scenarios upon which
various techniques are useful.

4.3 – Statistical Significance

Often, it is desirable to examine quantitative problems using a statistical lens, that
is, to determine if some pattern is different than an expectation. For those less familiar
with statistical inference in GISci, we point the reader to the text by O’Sullivan and
Unwin (2010), which provides an introduction to these concepts. Spatial statistics often
rely on the concept of complete spatial randomness (CSR) as an *a priori* assumption for
assessing the statistical significance of observed spatial patterns (Cressie 1993). With
some types of spatial statistics (e.g., join counts, Cliff and Ord 1981) the distributions for
computing statistical tests are analytically derived. With other statistics, specifically most
spatially local measures, simulation procedures are used to generate test distributions,
making these statistics primarily exploratory (Boots 2002).

Random walks have been suggested as being to movement data what CSR is to
spatial data (Winter and Yin 2010). Two key methodological developments have
included random movement in their derivation: Brownian bridge home ranges (Horne et
al. 2007) and probabilistic time geography (Winter and Yin 2010). However, these two
examples represent essentially the same problem: defining a probability surface for
movement between two known locations in space-time. Authors of both methods concede
that random movement is inappropriate for modeling objects that move non-randomly,
but contend that it represents a necessary starting point.
The development of space-time statistics for movement is still in its infancy and lacks clear direction for future research. Some have taken alternative views on this problem, for example treating movement data as a bivariate time series using spatial coordinates as dependent variables (e.g., Jonsen et al. 2003). Others have looked at geographic space first, often ignoring the temporal component altogether (e.g., Casaer et al. 1999). Both approaches are limited as they do not consider movement as a dynamic process that is a function of both space and time. To adequately address the process of movement, novel statistical techniques must consider space and time simultaneously in their derivation. This will be challenging however, as inferential statistics are ill-suited to the multidimensional complexity of movement (Holly 1978).

4.4 – Emerging Trends in Quantitative Movement Analysis

Technological advances now facilitate real-time capture and analysis of movement data on both wildlife and humans. In wildlife applications, real-time data acquisition is providing opportunities for conservation and wildlife management. Dettki et al. (2004) implemented a real-time tracking system for moose in Sweden, where data on moose movements could be used to initiate the start-up and shut-down of forestry operations in seasonal moose ranges. This idea relates directly to recent work identifying the importance of timing in time geographic measures of space-time accessibility (Neutens et al. 2010, Delafontaine et al. 2011a). As the interface between wildlife and humans narrows, other potential applications exist for real-time tracking. Consider a problematic large carnivore (e.g., lion or bear) residing in a national park. Rather than relocating or exterminating this animal, a real-time tracking system could be used to
monitor the animal’s movements. Park managers could use this information to improve park safety and minimize human-animal conflicts through trail/site closures and surveillance efforts.

Further developments with real-time movement data will involve the creation of increasingly sophisticated models for predicting future movement locations. The space-time cone from time geography (see Figure 1a) provides only the boundary for future movement possibilities (e.g., O’Sullivan et al. 2000), factoring in the uneven distribution of future movement possibilities (e.g., Winter 2009) provides more useful information for prediction. Future movement possibilities can be linked to contextual factors such as obstacles (Prager 2007), underlying movement cost surfaces (Miller and Bridwell 2009), and object kinetics (Kuijpers et al. 2011). Further developments towards probabilistically predicting future movements based on contextual factors will provide researchers and analysts with powerful tools for linking real-time movement data with other data sources.

With human movement data a new field that is gaining momentum focuses on leveraging real-time location data in everyday applications: location based services (Raper et al. 2007). Location based services have developed coincidentally with the availability of location-aware devices (e.g., GPS enabled cell-phones and handheld devices), which are now integral to people’s daily routines (Kumar and Stokkeland 2003). However, given the revealing nature of personal movement data, concerns over the privacy and ownership rights of personal movement information continue to surface (e.g., Dobson and Fischer 2003). With location based services, the fundamental goal is to tailor individual applications, services, and marketing to a user’s real-time location (Raper et al. 2007). For example, methods for predicting future movements based on
contextual factors, when applied in a real-time application, could provide increased functionality and improve user experiences with location based services. As methods for analyzing real-time movement data emerge, their development in conjunction with applications from location based services should be conducted in order to facilitate their adoption in this field.

With the development of technologies for acquiring movement data, the ability to capture finely grained movement data has increased substantially. Opportunities exist for investigating properties of movement previously not feasible with coarser grained movement data. For example, investigating velocities, accelerations, and the role of momentum in moving objects is an area of opportunity. Current research is developing methods for incorporating physical kinetics (based on object velocity and acceleration) into the calculation of time geography volumes, such as those from Figure 1 (Kuijpers et al. 2011). Another avenue for future work is the development of a probabilistic time geographic framework, such as by Winter (2009), that considers the influence of kinetics into the calculation of future movement probabilities.

Methods for investigating interactions between individuals in groups of moving objects continue to develop, but remain limited in overall scope and sophistication. Laube et al. (2005)’s relative motion concept can identify trendsetters, but uses only movement azimuth in its derivation. Others have developed other ways to identify specific types of interactions between moving individuals (e.g., Andersson et al. 2008; Buchin et al. 2010). As our ability to characterize these patterns grows, it may be more useful to investigate methods for quantifying the strength of interactions that occur in movement databases. That is, can we measure how interactive are the movements of two individuals. The work
of Shirabe (2006) provides a necessary starting point for this research which could be further investigated in light of this problem. Further, it may be necessary to examine outside factors influencing the levels of interaction between individuals (e.g., barriers and obstacles represented as lines/polylines, Noyon et al. 2007). Subsequently, how to accommodate other data sources into models for measuring individual level interactions in movement data remains an open research problem.

With time geography, Hägerstrand provided a theoretical context for looking at the constraints of object movement. Contemporary geographers continue to expand on time geographic concepts incorporating a range of ideas into time geographic theory (e.g., Winter 2009, Miller and Bridwell 2009, Delafontaine et al. 2011b). As discussed by Lenntorp (1999), Hägerstrand’s time geography represents a set of conceptual and methodological building blocks for use in analyzing and understanding movement as a process. As the quantitative toolkit for analyzing movement continues to grow and develop, those methods including theory and ideas from time geography in their derivation will have increased value in a broader range of applications.

Other theoretical frameworks have also been successfully implemented in movement research. For example, the idea that movement is motivated by an underlying field (e.g., Brillinger et al. 2001) suggests that forces of attraction and repulsion may influence movements. Such points of attraction, for example in wildlife, may be used to investigate central place foraging theory (Orians and Pearson 1979). Markovian models have also been used to demonstrate how movement operates as a diffusion process (e.g., Skellum 1951). Diffusion, originally used to describe random dispersal of organisms, can also be related to crowd dynamics in humans (Batty et al. 2003). The use of theoretical
constructs in quantitative methods, such as the aforementioned examples, demonstrates thoughtful development of ideas that in the end are easier to interpret for both the reader and analyst.

It has been suggested that movement methods must consider the “geography behind trajectories” (Bogorny et al. 2009) in order to understand the geographic processes affecting observed movement patterns. Movement analysis is no longer limited by available data, but rather by the tools required to manage and analyze movement databases in more efficient and sophisticated ways (Miller 2010). Thus, the continued development of methods capable of integrating increasingly large and complex movement databases with available spatial and temporal layers is warranted. With such analysis, the goal is to identify relationships between movement patterns and underlying spatial and/or temporal variables. Data mining work is beginning to enrich movement data with underlying geographic datasets (Alvares et al. 2007, Bogorny et al. 2009).

Quantitative methods for movement data must be further developed to consider underlying geographic variables in order for movement to be understood as a function of the environment. Similarly, novel movement datasets are emerging where attribute data are recorded along with spatial and temporal records (e.g., $<ID, S, T, A>$, where $A$ represents some attribute data). For example, wildlife tracking systems are being equipped with devices, such as cameras (Hunter et al. 2005), that simultaneously record information alongside movement fixes. The inclusion of attributes with movement fixes can be termed marked movement data, comparable to the term marked point pattern in the spatial statistics literature (Cressie 1993). Inclusion of attributes (numerical or categorical) alongside spatial locations in movement data represents an area of
opportunity for advanced analysis in the movement-attribute space, as existing methods are not designed for marked movement data.

6 – Conclusions

Novel movement datasets are not only becoming readily available they are changing how data on movement processes are captured. Traditionally, movement data have been collected as samples taken at coarse temporal granularities. Coarsely collected movement data represents movement discretely and with considerable uncertainty between sampled points. More recently, movement data are being collected at extremely fine temporal granularities, such as 5 fixes/second with athletes. Finely grained movement data represents a (near) continuous form of movement data which contains minimal uncertainty in space-time location. Not only are existing methods ill-suited for finely grained movement data, but the types of questions being asked must also be revisited to consider that uncertainty between consecutive fixes is negligible.

Within GIS data formats, there is a clear lack of appropriate structures for handling movement data. Those interested in purely visualizing movement data have circumvented these problems by generating independent platforms for visualizations (Andrienko et al. 2005). However, the development of quantitative methods is still hindered by difficulties representing the temporal domain within GIS. The development of geospatial data formats exclusively for movement data will invigorate future research into quantitative methods for movement.

There is a clear need for novel quantitative methods for extracting information and generating knowledge from ever-expanding movement datasets (Wolfer et al. 2001,
Laube et al. 2007). Most existing methods can be classified as data mining algorithms, which are used to identify and categorize trends in movement databases, based on some *a priori* notion about movement. Emerging problems investigate more complex patterns and relationships contained in movement datasets, such as the identification of flocking behavior (Benkert et al. 2008). Methods that are able to quantify interactions between individuals (Laube et al. 2005), and with environmental variables (Patterson et al. 2009) in movement databases will be increasingly relevant in more sophisticated movement analyses. Movement models capable of quantifying relationships between moving objects and dynamic features in the environment (e.g., traffic conditions) are justified in order to measure the significance of events or changes on object movement.

**Acknowledgements**

The authors would like to thank the referees for their constructive remarks on earlier versions of this manuscript. Their comments greatly improved the presentation of this work. Support for this work was obtained from the Natural Sciences and Engineering Research Council of Canada, and GEOIDE through the Government of Canada’s Networks for Centres of Excellence program.
7 - References


Forer, P., 1998. Geometric approaches to the nexus of time, space, and microprocess: Implementing a practical model for mundane socio-spatial systems. In:


Righton, D. and Mills, C. 2006. Application of GIS to investigate the use of space in coral reef fish: a comparison of territorial behaviour in two Red Sea


<table>
<thead>
<tr>
<th>Description</th>
<th>Term</th>
<th>Synonymous terms (with selected references)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A single record of object movement (of the form &lt;ID, S, T&gt;).</td>
<td>Movement Fix (Mt)</td>
<td>point, observation, relocation</td>
</tr>
<tr>
<td>A sequence of ordered records in time depicting the movement of a single object.</td>
<td>Movement Path (Mp)</td>
<td>space-time path (Hägerstrand 1970), trip-chain (Kondo and Kitamura 1987), geospatial lifeline (Mark 1998), trajectory, trace, track</td>
</tr>
<tr>
<td>A collection of records depicting the movements of many objects or the same object at different occasions, potentially including attribute information.</td>
<td>Movement Database</td>
<td>moving objects database (Güting and Schneider 2005)</td>
</tr>
</tbody>
</table>
Table 2: Parameters extractable from movement data sorted by dimension. After Table 1 from Dodge et al. (2008).

<table>
<thead>
<tr>
<th></th>
<th>Primitive</th>
<th>Primary Derivatives</th>
<th>Secondary Derivatives</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spatial</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(x, y)</td>
<td>Position</td>
<td>Distance</td>
<td>Spatial distribution</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Direction</td>
<td>Change of direction</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spatial extent</td>
<td>Sinuosity</td>
</tr>
<tr>
<td><strong>Temporal</strong></td>
<td>Instance</td>
<td>Duration</td>
<td>Temporal distribution</td>
</tr>
<tr>
<td>(t)</td>
<td>Interval</td>
<td>Travel time</td>
<td>Change of duration</td>
</tr>
<tr>
<td><strong>Spatio-temporal</strong></td>
<td></td>
<td>Speed</td>
<td>Acceleration</td>
</tr>
<tr>
<td>(x, y, t)</td>
<td>—</td>
<td>Velocity</td>
<td>Approaching rate</td>
</tr>
</tbody>
</table>
Figure Captions

Figure 1: Volumes used in Hägerstrand’s time geography: a) space-time cone, b) space-time prism, c) potential path area, and d) path bundling.
Figure 2: Comparison between definitions of a) flocks, and b) convoys. A flock requires objects be contained in a circle of radius $-r$, while a convoy is defined as those objects that are *density connected* at distance $-d$. Both methods require that objects be included in the group over a minimum time interval $-\tau$. 

![Diagram of flock and convoy definitions](image-url)
Figure 3: Four analysis levels for movement data: *instantaneous*, *interval*, *episodal*, and *global*. After Figure 2 from Laube et al. (2007).