Exploratory Visualization of Collective Mobile Objects Data Using Temporal Granularity and Spatial Similarity

Tetsuo Kobayashi and Harvey Miller

Abstract Recent advances in location-aware technologies have produced vast amount of individual-based movement data, overwhelming the capacity of traditional spatial analytical methods. There are growing opportunities for discovering unexpected patterns, trends and relationships that are hidden in massive mobile objects data. However, a lingering challenge is extracting meaningful information from data on multiple mobile objects due to the visual complexity of these patterns even for a modest collection of mobile objects. This chapter describes visualization environments based on temporal granularity, and spatial and/or attribute similarity measures for exploring collective mobile objects data. Reconstructing trajectories at user-defined levels of temporal granularity allows exploration at different levels of movement generality. At a given level of generality, individual trajectories can be combined into synthetic summary trajectories or classified into groups based on locational and/or attribute similarity. A visualization environment based on the space-time cube concept exploits these functionalities to create a user-interactive toolkit for exploring mobile objects data. A case study using wild chicken movement data demonstrates the potential of the system to extract meaningful patterns from the otherwise difficult to comprehend collections of space-time trajectories.

Keywords Spatio-temporal knowledge discovery • Temporal granularity • Mobile objects • Data aggregation • Geovisualization

T. Kobayashi (✉)
Department of Geography, Florida State University, 600 W College Ave, Tallahassee, FL 32306, USA
e-mail: tkobayashi@fsu.edu

H. Miller
Department of Geography, University of Utah, 201 Presidents Cir, Salt Lake City, UT 84112, USA
e-mail: harvey.miller@geog.utah.edu

G. Cervone et al. (eds.), *Data Mining for Geoinformatics: Methods and Applications*,
DOI 10.1007/978-1-4614-7669-6_7,

1 Introduction

Location-aware technologies (LATs) such those based on the Global Positioning System (GPS) or radio-frequency identification (RFID) chips have greatly enhanced capabilities for collecting data about mobile objects. LATs connected to location-based services (LBS) embedded in cellular telephones and other clients allow unprecedented access to individual mobility patterns across a wide range of domains (Brimicombe and Li 2006; Li and Longley 2006). GPS and RFID devices are increasingly connected to vehicles and objects in fleet management and logistics, generating fine-grained data on movements of these entities within supply chains (see Roberti 2003). Researchers in ecology and biology are also using LATs to track movements of animals, creating new insights into territoriality and ecosystem dynamics (e.g., Wentz et al. 2003; Turchin 1998)

The prevalence of LATs is generating a vast amount of mobile objects data that are overwhelming the capabilities of traditional spatial analytical methods. A major challenge in GIScience is to develop representation and analysis techniques that can handle spatio-temporal and mobile objects data (Laube et al. 2005). A related research challenge is developing methods to explore, analyze and understand the motions of collections of mobile objects over time. Analyzing one or a small number of mobile objects is tractable, but making sense of the collective mobility patterns of even a modest number of objects is daunting due to the visual complexity of the data involved (Shaw et al. 2008).

This paper describes a user-interactive visualization toolkit for summarizing and exploring mobile objects data based on spatial similarity among object trajectories at different levels of temporal granularity. Reconstructing trajectories at user-defined levels of temporal granularity allows exploration of the mobile objects at different levels of movement generality. At a given level of granularity, the user can apply similarity measures for aggregating or grouping trajectories based on location or spatial similarity. To maximize user-interactivity, the measures are computationally scalable to facilitate rapid calculation even on modest computational platforms. The similarity measures are also dimensionless and semantically-clear to facilitate easy interpretation. A visualization toolkit based on the space-time cube concept exploits these functionalities to create a user-interactive environment for exploring mobile objects data. A case study using wild chicken movement data demonstrates the functionality of the toolkit for extracting general patterns from an otherwise indiscernible collection of trajectories.

The visualization environment described in this paper intends to provide a user-friendly toolkit for scientists who are primarily concerned with data corresponding to objects moving through geographic space such as people, vehicles and animals. Therefore, we designed the methods and toolkit in this paper for objects that exhibit potentially continuous motion through space densely with respect to time. The temporal granularity and similarity aggregation methods are not designed for objects that exhibit discontinuous change at discrete moments in time. Therefore, other event or change data such as financial transactions or phone calls, while often referenced in time and/or space, are not appropriate.

The next section of this paper provides the background behind the concepts and methods used in this research. Following this, the methodology section explains the temporal granularity, similarity functions, and other visual and summarization functionalities of the visualization toolkit. Section 4 provides a case study that illustrates the methods using a space-time visualization toolkit. The final section summarizes the contributions of this research and suggests topics for further investigation.

2 Background

2.1 GIS and Mobile Objects

Since change with respect to time is common in most natural and human phenomena, incorporating time and change in geographic information systems has been a critical research frontier since the 1980s (see Langran 1992). Possible approaches include temporal "snapshots" that update the database at regular intervals, event-based approaches that update only the relevant portion of the data when a change occurs, and maintaining semantic, spatial and temporal dimensions in separate but linked domains (see Peuquet and Duan 1995; Yuan 2001; Worboys and Duckham 2004; Hornsby and Cole 2007).

Mobile objects data is an important special case of this general problem since these change their geometry near-continuous with respect to time. Explicitly updating of the geometry of a moving object is too expensive with respect to computational effort and storage requirements. Instead, one must accept some level of sampling error due to the finite and discrete updating of continuously changing objects and represent this error within the database (Sistla et al. 1998). Mobile objects also imply unique semantics and therefore a need for specialized query languages and analytical techniques (Andrienko et al. 2008). The field of mobile objects databases has emerged to handle these unique requirements.

Another set of techniques for understanding mobility data derives from the field of time geography and efforts to build GIS and other analytical tools based on its basic concepts. Time geography is based on the notion that the events that comprise an individual's or object's existence have spatial and temporal dimensions that are difficult to untangle in a meaningful way (Hagerstrand 1970). While only a conceptual framework traditionally, in recent years the applicability of time geography has been enhanced through the development of analytical and computational tools linked to GIS software (e.g., Kwan 2000; Miller 1991, 2005; Yu and Shaw 2008). However, most of these efforts address only a single or small number of mobile object trajectories due to time geography's bias towards the individual rather than collective behavior, as well as a lack of tools for handling collections of trajectories. For example, although Kwan (2000) develops interactive tools for visualizing mobile objects data, there are no capabilities for summarizing or aggregating these data, meaning that it is difficult to scale these applications to collections of trajectories without visual confusion. Shaw et al. (2008) addresses this issue by using clustering techniques to extract representational summary paths from trajectory collections.

2.2 Mobility Mining

The problem of extracting meaningful information from large databases is not unique to mobile objects data and GIS-based time geography. Knowledge Discovery from Databases (KDD) is the attempt to extract novel patterns hidden in massive digital databases through efficient computational techniques. The objective is to generate unexpected and interesting hypotheses that can be investigated further using standard inferential and confirmatory techniques. Geographic Knowledge Discovery (GKD) is a subset of KDD that attempts to discover novel spatio-temporal patterns in massive digital geographic datasets using scalable geocomputational techniques. GKD techniques exploit the unique characteristics of geographic data such as spatial dependency and heterogeneity. In addition, GKD tools can handle complex spatial properties such as the size and shape of geographic objects, and relationships among objects such as distance, direction and connectivity (Han et al. 2001; Miller and Han 2009). As the size and complexity of geospatial data increases, leveraging geocomputational techniques with geovisualization is essential to help manage the GKD process and interpret its results (Andrienko and Andrienko 2008).

Mobile objects data also creates unique challenges for the knowledge discovery process. Andrienko and Andrienko (2008) envision a specialized knowledge discovery process for these data called *mobility mining*. The mobility mining process involves three major steps:

1. *Trajectory reconstruction.* This involves processing the raw stream of mobility data to obtain the individual object trajectories. It also involves methods for efficient storage and access of these trajectories.
2. *Knowledge extraction*. This involves the application of spatio-temporal and mobile objects data mining methods to discover novel and useful information in these data. Possible patterns include *clusters* or groups of similar trajectories, *frequent patterns* reflecting repeatedly followed paths or subpaths and *classifications* based on behavioral rules extracted from the trajectories (also see Dodge et al. 2008).
3. *Knowledge delivery.* Extracted patterns are seldom direct knowledge; rather, these patterns must be evaluated based on their interestingness, interpreted relative to pertinent background knowledge and communicated in a manner appropriate for improving policy and decision-making in real-world applications.

Our main concern in this research is the rapid summarization of data as a first step in the knowledge extraction process. A well-known technique in online analytical processing (OLAP) technique is the *data cube*. The data cube is an operator that allows users to generate all possible cross-tabulations of the data at different levels of aggregation to provide synoptic summaries of the database (see Gray et al. 1997; Han and Kamber 2006). Shekhar et al. (2001) extended the data cube to the *map cube* that can handle the geographic components of the data and visualize them in concert with the cross-tabs and aggregations. The *traffic cube* is a further extension for handling spatio-temporal traffic data (see Lu et al. 2009; Shekhar et al. 2001, 2002). However, these methods require a fixed geography and spatial aggregation scheme and cannot be applied directly to trajectory summarization.

In addition to database aggregation methods, data visualization techniques enable simple and intuitive interactions of mobile object data and humans. The objective is to find interesting patterns, trends and relationships especially in mobile objects datasets, supporting knowledge construction about mobility behavior (Miller and Han 2009). Dynamic visual exploration is useful in understanding the structure of the dataset, raising questions about movement patterns, and facilitating identification of meaningful combinations of variables in further map representation and analysis (Wood and Dykes 2008). Interactive visualization methods associated with visualization software environment have been proposed to enhance the quality of pattern detection. For example, GeoTime is a three-dimensional visualization environment designed to visualize and analyze trajectories of mobile objects (Kapler and Wright 2005; Kraak 2003). Another widely used method for visual data exploration is trajectory aggregation. Data mining methods such as clustering support visual detection of clusters of mobile objects for data aggregation (Andrienko et al. 2009; Rinzivillo et al. 2008; Schreck et al. 2008). Although these studies successfully illustrate the importance of analysing spatio-temporal dynamics within a visualization environment, they heavily rely on locational and temporal information only. Few studies have explored the mobile object patterns from the attribute domain (Skupin 2008; Kraak and Huisman 2009).

An exploratory visualization technique designed specifically for mobile objects data is the *space-time cube* (Kraak 2003). The space-time cube visualizes spatio-temporal data in a three-dimensional environment that the user can manipulate through rotating, projecting, scaling and other visual browsing techniques (Kraak 2003) (Fig. 1). In addition, Leonardi et al. (2010) developed the *T-warehouse* for data warehouses system designed for trajectory data.

2.3 Data Aggregation and Similarity Measures

As noted above, a barrier to meaningful visualization of mobility databases is difficulty in extracting meaningful patterns from mobile objects data. *Data aggregation techniques* are methods for reducing the size of data to extract general patterns (Andrienko and Andrienko 2008). Several researchers have proposed time-based aggregation to summarize and analyze mobile objects data. For example, Hornsby and Egenhofer (2002) developed a framework that enables space-time

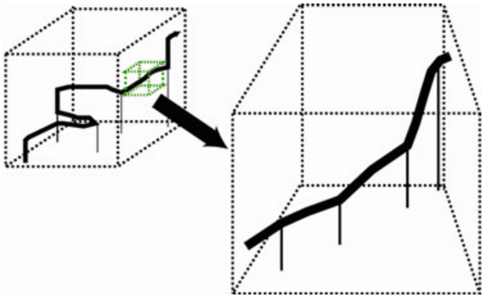


Fig. 1 The space-time cube concept

queries in multiple time granularities for space-time paths and prisms. In addition, there are some efforts to combine time geographic concepts and data summarization methods such as aggregation and clustering. These software tools visualize mobile object trajectories in two spatial dimensions and time, and provide capabilities to group trajectories based on location during a time period of interest (Pfoser and Theodoridis 2003; Shaw et al. 2008; Kapler and Wright 2005).

An emerging data aggregation technique for mobility data is *similarity measures*. Similarity measures can be used to analyze whether different mobile objects exhibit correspondence in terms of a metric distance function such as Euclidean distance (Sinha and Mark 2005; Yanagisawa et al. 2003), the Hausdorff distance measure for two point sets (Huttenlocher et al. 1993; Shao et al. 2010), the Frèchet distance for polygonal curve similarity (Eiter and Mannila 1994), and longest common subsequence (LCSS) for measuring similarity in time-series data (Vlachos et al. 2002). However, similarity measures alone do not allow the analyst to explore similarity at multiple scales (e.g., Laube et al. 2005). In addition, similarity measures can be computationally complex, although progress has been made with respect to scalable heuristics (Andrienko et al. 2009; Shao et al. 2010; Sinha and Mark 2005).

Another trend in similarity measure is trajectory descriptors. Trajectory descriptors are metrics of mobility physical characteristics such as location, direction and speed. These measures can serve as a basis for aggregating or grouping individual paths for summarization, improving the clarity of the visualization (Laube et al. 2005; Sinha and Mark 2005). Trajectory descriptors can be calculated at an individual sample location and can be extended into interval and/or global scales (Dodge et al. 2008). However, most studies focus on spatial and temporal domain; it is rare to examine dynamics within the attribute domain, that is, the evolution of non-locational properties over time such as the trajectory geometry and other physical movement parameters. This is also rarely linked with the growing area of geographic data mining and knowledge discovery (Skupin 2008; Skupin and Hagelman 2005).

The toolkit described in this paper combines similarity techniques with user-defined temporal granularity parameters to facilitate exploratory trajectory aggregation at varying levels of movement generality. Furthermore, the techniques developed in this visualization toolkit are computationally efficient and can be scaled to large databases and embedded in other exploratory techniques and processes. In addition, this paper describes a user-interactive visualization environment to summarize and explore mobile objects data based on the movement attributes of mobile object trajectories at different levels of temporal granularity. The next section of this paper discusses these time aggregation and similarity measure techniques, and the visualization environment that implements these techniques.

3 Methodology

3.1 Overview

This chapter develops time aggregation methods and similarity measures to enhance the discovery of multi-scale patterns in mobile objects data. There are several steps to analyze the mobile objects data within the interactive visualization tool proposed in this research (see Fig. 2). First, all the data are stored in a database in order to be extracted later as queries. Second, time aggregation methods allow the user to determine a time range of interest and temporal granularity within the selected time range to reconstruct individual trajectories at different levels of movement generality. Third, given these reconstructed trajectories, the user can apply similarity measures to aggregate individual trajectories based on location or attribute to aggregate multiple trajectories into synthetic trajectories that reflect collective movement patterns. This process can be repeated until the user of the toolkit finds meaningful patterns. We embed these techniques within a space-time cube environment that allow visual exploration and statistical summaries of the aggregated and grouped mobile objects data.

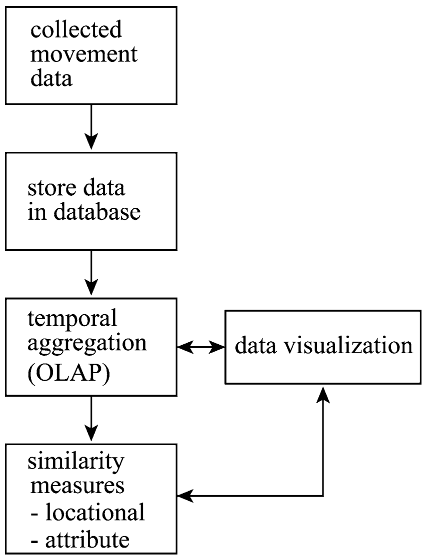


Fig. 2 Flowchart of the analysis

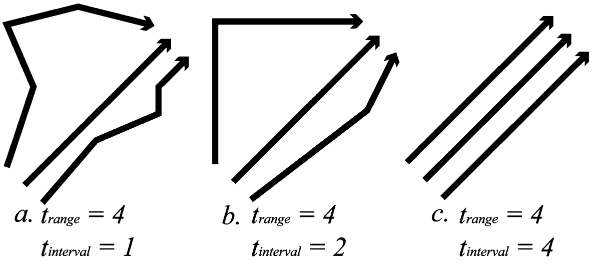


Fig. 3 Example of time range and time interval parameters

3.2 Time Granularity and Trajectory Reconstruction

Temporal granularity is a critical parameter for visual data exploration as well as data mining and statistical analysis since it can cause substantial difference in the results of visualization and analysis (Hornsby 2001). Visualization with coarse time granularity is more appropriate to explore broad scale movement while visualization with refined time granularity is more suitable for detailed movement of the mobile objects (Hornsby and Egenhofer 2002).

Two parameters for determining time granularity when reconstructing mobile object trajectories are the *time range* and *time interval*. Time range is the time period queried from the database. For example, if the user wants to visualize parts of trajectories at time between 10:00 and 11:00, ‘1 hour’ is the time range. On the other hand, time interval is the granularity within the time range; the minimum time unit that divides time range equally. For example, if the time range is 1 hour and the time interval is 10 minutes, the number of time stamp is six. The reconstructed trajectories reflect the choices of range and interval. To illustrate, assume trajectories from the database such as the ones as illustrated by Fig. 3a. As the time interval increases, three trajectories become more similar as in Fig. 3b, and exactly the same as shown in Fig. 3c.

Since LATs often record trajectory data using independent sampling rates, we normalize the trajectory data to common sampling times using simple temporal resampling rules. We map recorded locations and times to the interval that includes that sample. If there is more than one sampled point within the interval, we choose the first one in sequence. If no sample point falls within an interval, we interpolate the location and time based on its neighboring intervals. This resampling rule is efficient and scalable; however, note that choosing longer time interval may cause distortion of sampled locations of trajectories because the algorithm proposed in this research chooses the first time-stamped location within the chosen time interval. This resampling process enables comparison of trajectories recorded at different temporal intervals and granularities.

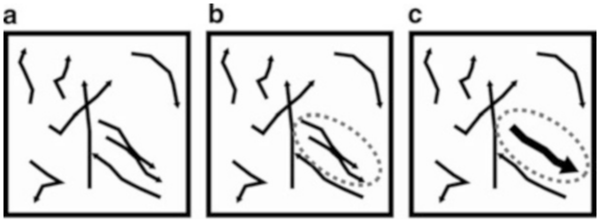
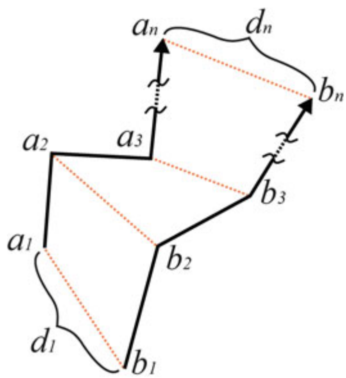


Fig. 4 Similarity aggregation: (**a**) individual trajectories; (**b**) detection of a cluster based on locational similarity; (**c**) summary trajectory

Fig. 5 Calculating locational similarity



3.3 Similarity Measures

Locational Similarity

This function measures the similarity between trajectories based on their spatial footprints, allowing users to aggregate trajectories that are spatially proximal given the selected time interval and range. This measure is useful in finding where and when mobile objects are moving together in a sequence of time. Urban commuting behavior, normal crowd flow and animal flocking are example movement patterns that can exhibit locational similarity. Figure 4 illustrates the general process.

An efficient method to measure locational similarity is to calculate the Euclidean distance between two locations of mobile object trajectories at specified time intervals (Steiner et al. 2000). A Euclidean distance of zero indicates that two trajectories visit the same locations in space-time. Trajectories that share the same spatial locations but at different times will have a higher locational similarity score, as will trajectories that diverge in space, even if they share the same origin, destination or some intermediate locations (Fig. 5).

If trajectories have a high degree of locational similarity, we can meaningfully aggregate these into a trajectory that summarizes those locations. A simple and

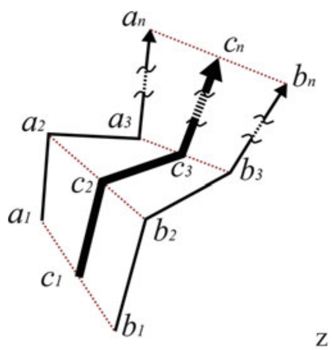


Fig. 6 Aggregating trajectories with locational similarity by vector averaging

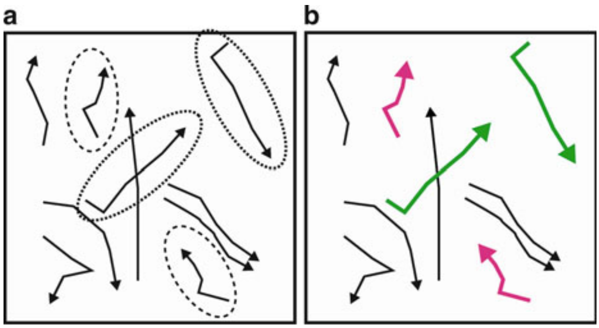


Fig. 7 Attribute similar trajectories: (**a**) detection; (**b**) identification

tractable method is to treat each polyline segment as a vector and finding the average of the corresponding vectors. Figure 6 illustrates the process for the two trajectory case for clarity.

While the Euclidean distance measure is straightforward, it suffers from sensitivity to outliers. Other possible distance measures include the Hausdorff and Fréchet distances. The Hausdorff distance is the maximum of the minimum distances between two curves; however, it can be misleading since it does not consider any temporal sequencing in the curves. The Fréchet distance captures the sequences within each curve (see Alt et al. 2003). We use the Euclidean distance measure for simplicity and scalability. However, our methods are not limited to Euclidean distances, and continuing development of the toolkit could include other distance measures for comparison purposes.

Attribute Similarity

In contrast to locational similarity, attribute similarity concerns intrinsic attribute properties of the trajectories regardless of their location or orientation in space. These measures can be used to categorize trajectories into groups based on similar attribute properties (see Fig. 7).

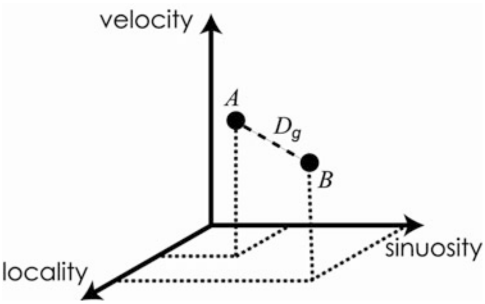


Fig. 8 Attribute similarity between two space-time paths based on three dimensions

The attribute similarity functions in the visualization toolkit extract five measurable attributes of attribute similarity. These are *sinuosity*, *direction*, *velocity*, *locality* and *spatial range* (described in detail below). The indices map the trajectory to a point in a multidimensional space (see Wentz 2000). The Euclidean distance within this space represents the similarity or dissimilarity of trajectories based on attribute. Figure 8 illustrates this in three-dimensional space for clarity. $D_g = 0$ indicates that two trajectories are exactly the same in terms of attribute, with increasing positive values indicating greater attribute dissimilarity. The user sets a maximum D_g values as a threshold to detect groups of similar trajectories for aggregation.

Although reducing complex attribute similarity properties to a single point in multidimensional space results in information loss, it creates a simple measure for efficient clustering, as well as input into other data mining techniques. This places a burden on the user to explore a wide range of similarity thresholds and assess the resulting summary patterns. It is therefore critical that the implemented system have a high degree of user interactivity.

The subsections below describe the five attribute similarity indices. The five indices are not independent, nor do they exhaust all aspects of attribute similarity; see Andrienko et al. (2008), Dodge et al. (2008), and Huang et al. (2008) for discussions of other attribute similarity measures. The indices proposed in this research are semantically clear properties that can be captured in an efficient, scalable manner and expressible as dimensionless metrics for ease of comparison. The user can apply all the indices simultaneously, or any subset depending on the nature of the data and the relevant questions to be explored.

- *Sinuosity*. Sinuosity measures the deviation of the trajectory from a straight line. It is the ratio of the total length of the trajectory and the Euclidean distance between the origin and destination:

$$Sinuosity = \frac{d_E}{d_p} \tag{1}$$

  where d_p is the total length of the trajectory and d_E is the Euclidean distance between the origin and the destination. This index varies between zero and one,

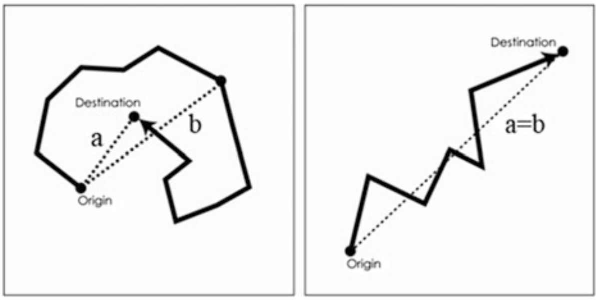


Fig. 9 Locality. Left side illustrates a low locality score; right side illustrates a high locality score

with one corresponding to a straight line and values closer to zero indicating a more sinuous path. The index is zero in the degenerate case of stationary behavior.

- *Direction*. This index captures the relative or egocentric direction of the trajectory:

$$Direction = \frac{\bar{D}}{180} \tag{2}$$

where $\bar{D}$ is the average egocentric direction of the line segments comprising the trajectory. Each line segment's egocentric direction is relative to the previous segment. Note the contrast with locational similarity: this considers directionality but from a cardinal perspective (e.g., two trajectories must travel in the same cardinal direction to have a low locational similarity value). In this index, two trajectories must have analogous tendencies with respect to turning directions to be similar. Since the egocentric direction ranges from $-180°$ (left hand side direction, or counter clockwise direction) to $180°$ (right hand side direction, or clockwise direction), the value of *Direction* ranges between -1 and $+1$.

- *Velocity*. Velocity indicates the relative speed of the object during the time period:

$$Velocity = \frac{\bar{V}}{V_{max}} \tag{3}$$

where $\bar{V}$ is the average velocity of a trajectory, and V_{max} is the maximum velocity in the data based on the used-defined temporal granularity parameters. Zero indicates stationary behaviour and one indicates matching the maximum velocity in the sample.

- *Locality*. Locality is the ratio between the distance between a trajectory's origin and its final destination and the distance between the origin and the farthest location in the trajectory (relative to the origin; see Fig. 9). This is a measure of the relative focus of the trajectory with respect to its initial and final location within the chosen temporal range.

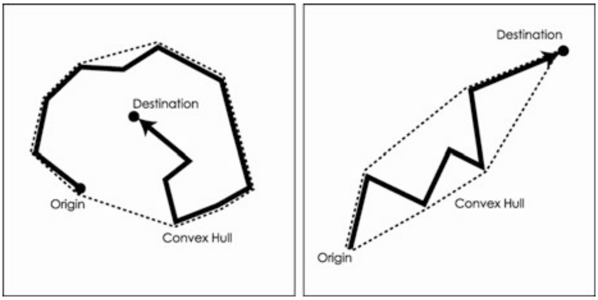


Fig. 10 Convex hull of a path

$$Locality = \frac{L_{OD}}{L_{OF}} \tag{4}$$

where L_{OD} is the Euclidean distance between the origin and the destination, and L_{OF} is the distance between the origin and the farthest recorded location from the origin. With respect to human movement behavior, trajectories with higher locality scores (closer to one) tend to be focused and single purpose, while trajectories with lower locality scores (closer to zero) tend to be more leisurely and/or multipurpose. For example, with respect to animal movement, low locality scores may indicate searching or foraging behavior. Using the origin as a basis for calculating locality may appear arbitrary: one could choose the destination instead. However, using the origin as a reference reflects the common practice in transportation science of characterizing trips based on their origin. It is also possible to use the origin and destination as joint references for a more complete depiction, but this would make the behavior of the index less transparent.

- *Spatial range*. Spatial range measures the relative spatial extent of the movement. It is the area of convex hull that contains a trajectory divided by the area of the convex hull that contains all the trajectories in the database:

$$Spatial\,Range = \frac{A_{path}}{A_{all}} \tag{5}$$

where A_{path} is the area of convex hull that contains individual trajectory (see Fig. 10) and A_{all} is the area of convex hull that contains all of the trajectories (see Fig. 11). A spatial range closer to zero indicates that the trajectory covers relatively little territory, while a spatial range closer to one indicates a more expansive territory for the object. A convex hull provides a relatively accurate measure of spatial range (relative to other measures such as the minimum bounding rectangle) with reasonable computational cost. The toolkit utilizes the Graham scan algorithm: this has the worse-case time complexity of O(n log n); this is better than quadratic and therefore scalable (Sedgewick 1990).

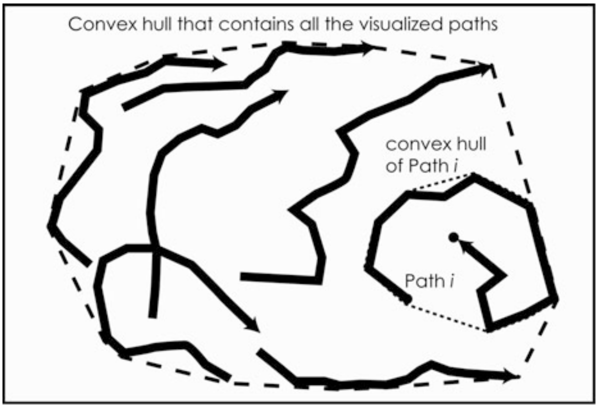


Fig. 11 Spatial range

Once the system calculates attribute similarity values, a next step is to generate groupings of trajectories based on these values. We look for natural groupings using an efficient density-based clustering method called DBSCAN (Ester et al. 1996). This is a spatial clustering method that looks for regions with sufficiently high density and forms clusters of arbitrary shapes. Although clustering methods designed explicitly for trajectory data are available (see Han et al. 2009), we chose DBSCAN due to its ability to incorporate the user-selected similarity indices discussed above as well as its scalability. DBSCAN also does not require designating the number of clusters *a priori.*

DBSCAN requires two input parameters: (i) an ε *-neighborhood* (expressed as a radius) for searching around each data point; (ii) *min-points* or the minimum number of data points required for in a neighborhood to be included. DBSCAN finds clusters by searching the ε -neighborhood of each data point, starting with an arbitrary point. If the ε -neighborhood of a point contains more than min-points, a new cluster is generated with that point designated as a *core object*. The algorithm iteratively adds data points that meet the search radius and density criteria to the core objects until no more points can be added. DBSCAN is efficient: it has the worst-case complexity of $O\ (n\ \log n)$ if a spatial index is used (Han and Kamber 2006; Han et al. 2009).

We customized DBSCAN in the following manner. First, the distances for finding neighboring points are the distances provided by the selected similarity functions. Second, we set min-points as one by default, indicating a cluster can be created from only two trajectories. Determining the ε -neighborhood for searching is more complex and can require trial-and-error exploration. To facilitate this, the visualization toolkit reports a set of statistical values for the data, including minimum and maximum values of the spatial coordinates in each dimension. A third modification is the inclusion of a maximum radius (*max-* ε) to limit the search around each point for scalability purposes. We set *max-* ε equal to twice the ε -neighborhood as a default, although this can be user-modified. This clustering algorithm is basically a combination of both DBSCAN and OPTICS (Han and Kamber 2006).

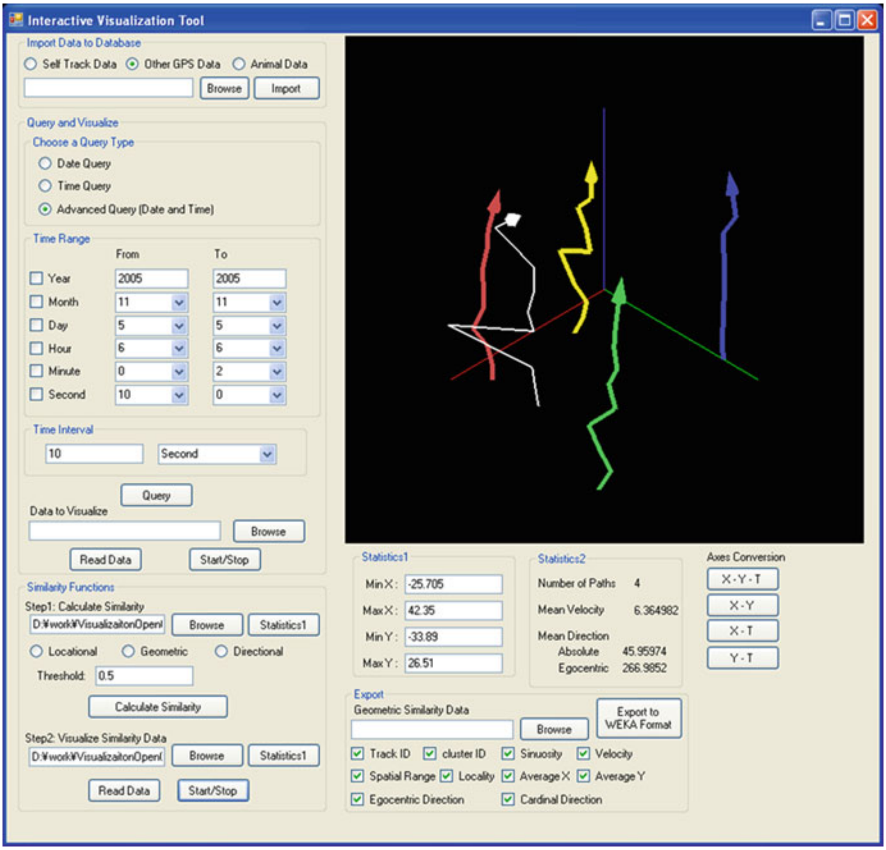


Fig. 12 Graphical user interface of the visualization tool

3.4 *Visualization Toolkit*

The visualization toolkit in this paper is based on the concept of the space-time cube (Kraak 2003). The toolkit encompasses three major functionalities: (i) individual trajectory reconstruction based on time granularity (time aggregation); (ii) trajectory aggregation and/or grouping based on trajectory similarity (similarity functions) and; (iii) data visualization, data summarization, and data export for further analysis. We developed the visualization toolkit using the C# programming language. In addition, Microsoft SQL Server 2005 provides the functionality for data storage and query support.

Figure 12 illustrates the main GUI. The user can visualize and explore these individual trajectories, aggregate or group trajectories based on their apparent similarity and extract summary statistical properties for the aggregated or clustered trajectories. The user can also visualize the trajectories in three dimensions (two-dimensional space and time), as well as project the trajectories into any two of the dimensions (*xy*, *xt* and *yt*). Arbitrary rotation of the visualization interface is also available. It is also possible to export data files containing the trajectory data and calculated parameters at any level of aggregation for import into other databases, data mining or statistical software. This allows the user to store detected patterns for further investigation, or as reference patterns for comparison.

4 Case Study

The aggregation methods proposed in this study can be applied to any mobile objects data where a sequence of time-stamped location stamps records the trajectory of each object. To illustrate the effectiveness of aggregation methods and other functionalities in the visualization toolkit proposed in this paper, we present results based on wild chicken tracking data in Thailand. Note that we use this data to illustrate the capabilities of the aggregation methods and functionalities of the visualization toolkit for extracting patterns from an otherwise mass of unintelligible trajectories. We do not intend to focus on the behavioral aspects of wild chickens and therefore will not offer potential hypotheses for the extracted patterns.

4.1 Mobility Data

The Human-Chicken Multi-relationship Research (HCMR) Project conducted a tracking analysis of wild chickens in Chiang Rai, Thailand using a Wireless Fidelity (WiFi) positioning system (Okabe et al. 2006). A small WiFi tag attached to a chicken's body records the location and time. The WiFi data tracking system consists of six devices, namely tags to stick to the chickens' legs, activator, receiver, Power over Ethernet (PoE), WiFi access point, and management engine. The weight of the tag is 35 g. Spatial resolution is 1 m and time resolution is 1 s. The fine time resolution of 1 s allows flexibility to analyze the data from detailed temporal scales to coarse temporal scales.

The study area is the 200 square-meter land under cultivation inside the Chiang Rai Livestock Research and Technology Transfer Center. Figure 13 illustrates the study setting, including the facilities and the locations of the chickens at one moment in time. There are eight concrete one-storied houses (H1 through H8) in the field: two of them are residential houses (H1 and H5) and the rest of them are empty houses. H2 is the preparation room for experimental appliances and H6 is the room for the data management engine that includes the software package that processes signals of location data sent via the WiFi access point and displays the locations of tags (Okabe et al. 2006). CH1 through CH3 are the locations of chicken houses. Since all the residents in two residential houses leave for agricultural work outside of

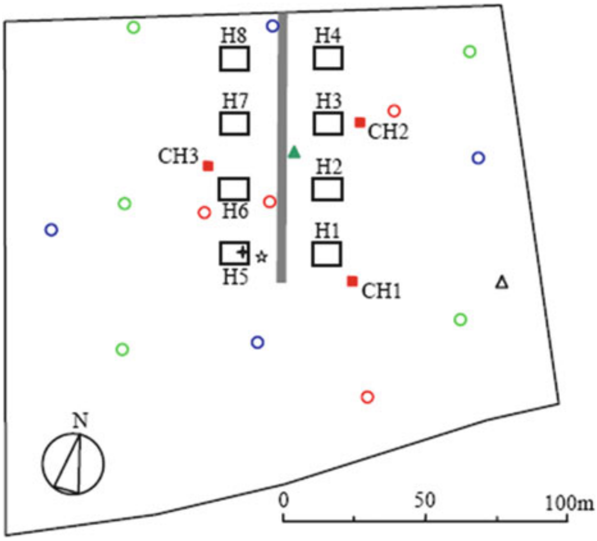


Fig. 13 Setting for wild chicken movement study (cited from Okabe et al. 2006)

the study field, chickens can move freely all over the study area. There are eighteen chickens in total (circle shape symbols in Fig. 13): there are three groups and each group consists of six chickens respectively. There are other symbols representing feeding sites (gray-colored triangle), trees (black, hollow triangle and star-shaped symbol), and the location of underfloor in the house H5 (cross-shaped symbol).

This research utilizes the movement data of 18 chickens with 2,979,359 time stamps from November 5th to November 8th 2005. Although the WiFi system uses x and y coordinates for locations, the location coordinates used in the system are independent from the geographic coordinates. The maximum spatial extent of the whole movement of the chickens can be expressed by maximum and minimum coordinates for both x and y coordinates. The minimum and maximum x coordinate are -83.79 and 80.39 respectively (164.18 in total for east–west extent), the minimum and maximum y coordinate is -53.71 and 49.42 (103.13 in north–south extent).

4.2 *Toolkit Functionality*

We now illustrate the toolkit functionality by showing results from querying the database at different levels of temporal granularity and aggregating the trajectories based on similarity at the specified granularity.

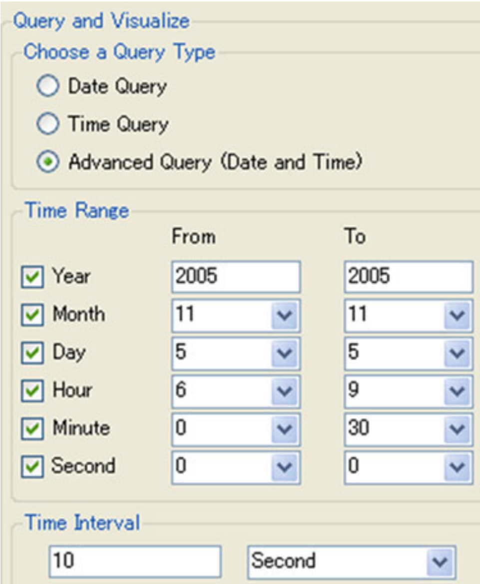


Fig. 14 Temporal aggregation using OLAP in the toolkit

Time Granularity and Trajectory Reconstruction

The proposed toolkit in this research enables the user to specify the time range and time interval of interest. Figure 14 shows the temporal aggregation GUI illustrating an OLAP query. The user first chooses the type of query from three options; date query, time query, or the advanced query that can specify both date and time. The user chooses the query type based on one's interest to extract a portion of trajectory data. The second parameter is the time range. This example is the case when the time range is between 6:00:00 and 9:30:00 on November 5th in 2005. The user can specify each temporal resolution from the drop down menus. The third and the last parameter is the time interval. This example shows the case of 10 s. The user again can choose the time resolution from the drop down menu. The options are years, days, hours, minutes, and seconds.

Figures 15 and 16 show the effects of time granularity on trajectory reconstruction. Figure 15 illustrates the reconstructed trajectory collection for the wild chicken data at three different time ranges on November 5th, 2005, with the time interval provided by the data (1 s). Figure 16 illustrates the reconstructed wild chicken trajectory collection at three time intervals for a fixed time range from 6:00 to 17:00 on November 5th, 2005. As Fig. 15 suggests, it is difficult to extract distinct patterns at the highest level of temporal granularity. Even with a relatively low time range (6:00–9:00), the trajectory collection is an undistinguished mass. This problem becomes more acute as the time range increases. Note that the map at the bottom of the visualization window shows that spatial extent of the whole movement also expands as the time range increases. Figure 16 indicates that changing the time

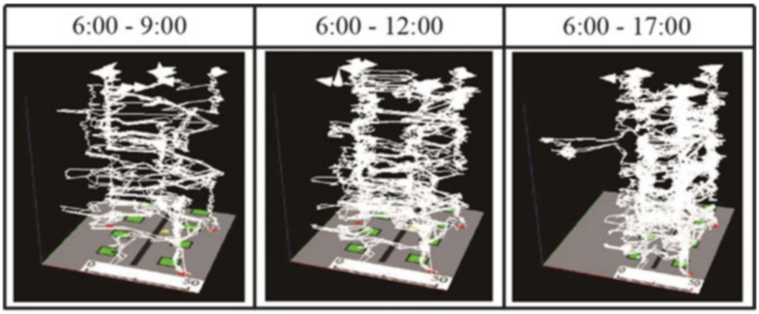


Fig. 15 Reconstructed trajectories at different time ranges

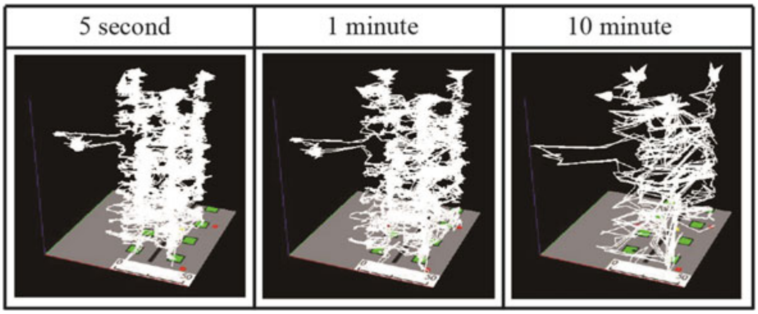


Fig. 16 Reconstructed trajectories at different time intervals

interval can mitigate this problem to a substantial degree: the trajectories are more generalized and patterns are more easily discernable as the time intervals become coarser. The visualization toolkit allows the user to visualize the trajectory collection at the time range of interest and interactively change the time interval until an appropriate granularity level is achieved for the data and questions at hand.

Locational Similarity

After the user has selected the time range and interval, the toolkit allows aggregation of similar trajectories to detect clearer patterns from the data. Figure 17 illustrates the effects of locational similarity-based trajectory on the visualized patterns at different time ranges for the wild chicken data in a three dimensional view. Figure 17 compares the unaggregated trajectories from Fig. 15 (top row in Fig. 15) with aggregated trajectories based on a strict locational similarity threshold of 5.0 (middle row) and a relaxed locational similarity threshold of 10.0 (bottom row). The top row once again illustrates the problem with unaggregated trajectories: it is difficult to discern any generalized patterns. In contrast, aggregation based

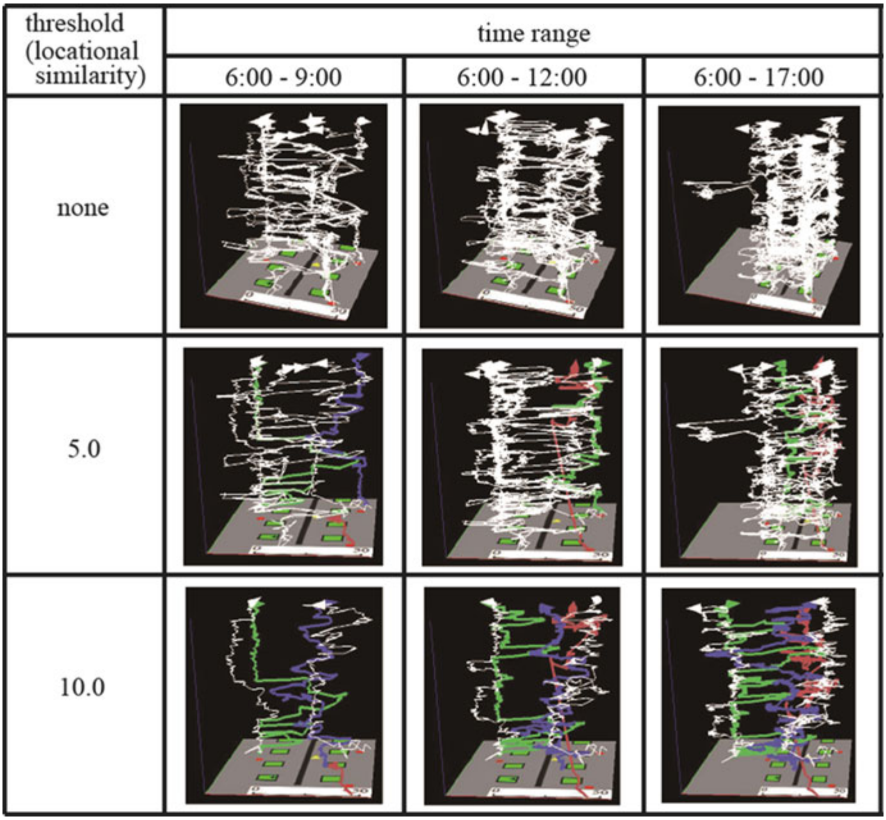


Fig. 17 Aggregated trajectories based on locational similarity

on locational similarity facilitates the detection of movement patterns. Summary paths (rendered blue, green and red) were extracted at both threshold levels. At the strict locational similarity threshold of 5.0, three summary paths were extracted for the time range 6:00–9:00, but only two paths for the longer time ranges of 6:00–12:00 and 6:00–17:00. In addition, these summary paths are occluded by the outliers (rendered in white) meaning that they represent a relatively small number of the sample trajectories. Three summary paths that are relatively stable across all three time ranges were extracted at the more relaxed threshold of 10.0. Also, these summary paths are easier to discern since the number of outliers is smaller. The appropriate value for this threshold must be determined by user-interactivity: the toolkit facilitates this process.

The toolkit also reports statistical data for the aggregated paths: this can help with user interpretation of the results. Table 1 provides some of the statistical data for the aggregate trajectories in Fig. 17. Based on the statistics in Table 1, directional values become close to zero as the time range increases in all three clusters regardless of

Table 1 Statistical information of summary trajectories in Fig. 17

Locational similarity threshold	Cluster color	Time range	Length	Mean velocity	Mean cardinal direction	Mean egocentric direction
5.0	Red	6:00–9:00	34.00	2.62	15.25	178.28
		6:00–12:00	475.00	3.44	1.19	8.63
		6:00–17:00	1273.25	2.91	0.42	2.22
	Green	6:00–9:00	520.45	1.45	1.38	17.48
		6:00–12:00	894.50	1.24	0.55	−144.40
		6:00–17:00	1388.52	1.05	0.54	1.53
	Blue[a]	6:00–9:00	422.89	0.78	1.11	−127.30
10.0	Red	6:00–9:00	34.00	2.62	15.25	178.28
		6:00–12:00	475.00	3.44	1.19	8.63
		6:00–17:00	1250.28	2.76	0.56	1.04
	Green	6:00–9:00	485.73	1.35	1.78	7.72
		6:00–12:00	945.98	1.31	0.04	7.96
		6:00–17:00	1658.74	1.26	0.31	6.83
	Blue	6:00–9:00	532.38	0.99	1.46	−20.03
		6:00–12:00	939.12	0.87	0.74	−26.49
		6:00–17:00	1706.96	0.86	0.10	−7.99

[a]No clusters occurred during the 6:00–12:00 and 6:00–17:00 time ranges

the locational similarity threshold. This implies the movements are not fixed in a certain direction. In addition, the mean cardinal direction is always close to zero degree in all cases, indicating that movement tend to direct all directions. Also, the mean velocity of red cluster increases as the time range increases from 6:00–9:00 to 6:00–12:00 while the mean velocity of other clusters decrease as the time range increases regardless of the locational similarity threshold.

To help users identify the numbers of detected clusters, it is useful to observe the change in the number of detected clusters in relation to time interval and time range. Figure 18 traces relationship between the change in the number of detected clusters of wild chicken data in three different time intervals and locational similarity threshold values. As is shown in the Fig. 18, number of clusters changes in a similar manner in all three time intervals indicating there are similar clusters detected regardless of the difference in time intervals. Figure 19 shows the relationship between the change in the number of detected clusters of wild chicken data in three different time ranges and locational similarity threshold values. The number of clusters varies with different time ranges although the trend in change of the number of clusters is similar in all three time ranges. Generally, the higher the locational similarity threshold, the more trajectories are likely to be included in fewer numbers of trajectories, resulting in detecting only one cluster when the threshold value is very large. However in this case, the number of clusters converges to either three or four indicating there are distinct differences in those three or four groups of movement.

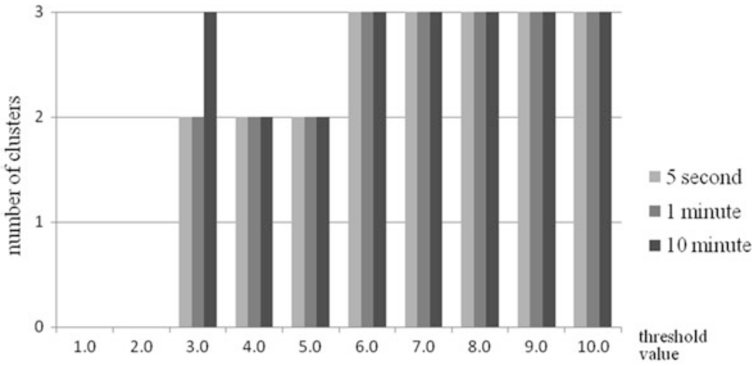


Fig. 18 Change in the number of detected clusters of wild chicken data by locational similarity measure with respect to time interval

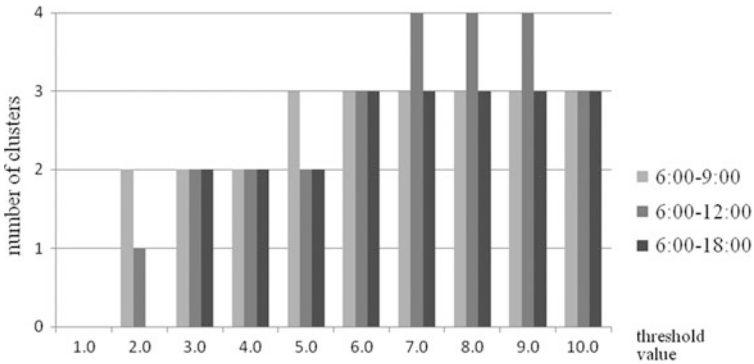


Fig. 19 Change in the number of detected clusters of wild chicken data by locational similarity measure with respect to time range

In addition to the three dimensional view, Fig. 20 clearly shows the spatial distribution of the movements of chickens. The aggregated trajectories with a relaxed locational similarity threshold of 10.0 detected the same three clusters from shorter time range (6:00–9:00) to the longest time range (6:00–17:00). This indicates three important findings. First, the locational similarity measure successfully detected three clusters that are reported in Okabe et al. (2006) that used the same data set. Okabe et al. (2006) reported that there are three main groups of chickens that behave as flocks for the entire study period. Second, the locational similarity measure detected the clusters at similar locations throughout the day, which implies the consistency in the movement of three groups of chickens. Chickens in this case tend to move as groups although there are some outliers (trajectories in white). Third, locations of all three detected clusters overlap or are close to the location where the food is (the triangle point in Fig. 14). The chickens did not move through wide areas of the study area but stayed close to where houses are located.

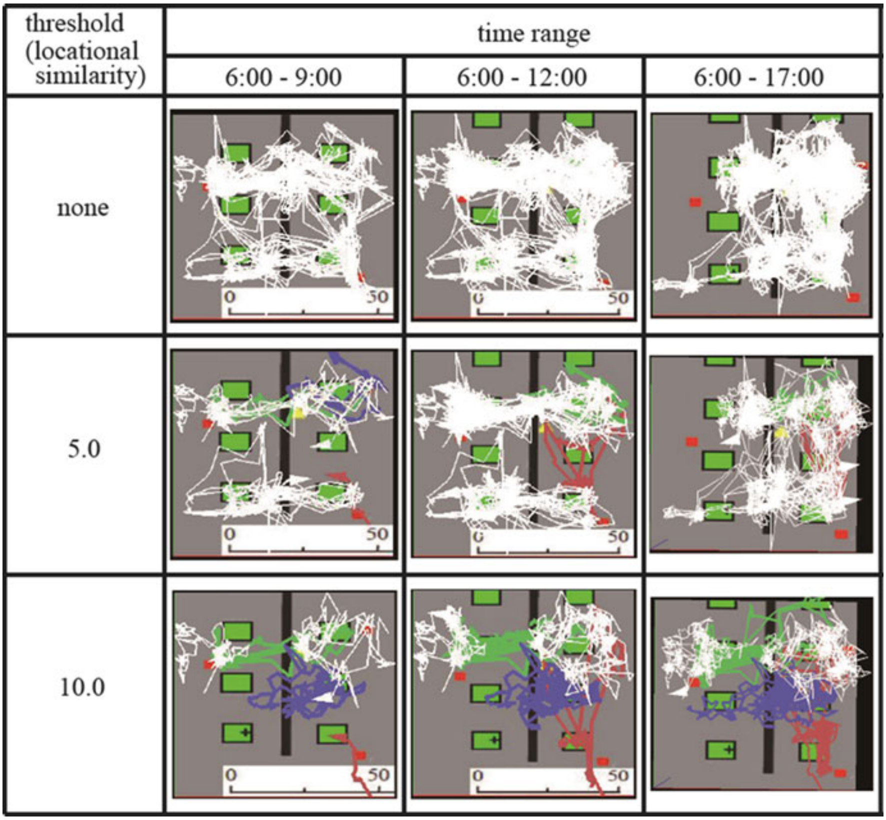


Fig. 20 Two dimensional view of aggregated trajectories based on locational similarity

Attribute Similarity

Figure 21 shows visualizations of the wild chicken trajectory collection at different attribute similarity thresholds and for different time ranges. The time interval is fixed at 10 s and all five attribute similarity functions are invoked. Comparable to locational similarity, a strict attribute similarity threshold (0.1) only detects a small number of trajectories for clustering, but a more generous threshold (0.5) identifies a greater number of candidates. Obvious patterns appear when attribute similarity function is applied compared with the visualization without any similarity functions (top row). In addition, similar patterns with respect to attributes tend to appear at similar locations: paths with similar attribute properties tend to occur in proximity, suggested coordinated movement behavior. It is also interesting that the locations of clusters detected in two time range, 12:00–13:00 and 16:00–17:00, are similar to each other. Chickens may move the same areas in different time ranges: this suggests repetitive movement patterns. Okabe et al. (2006) also reported that there are some

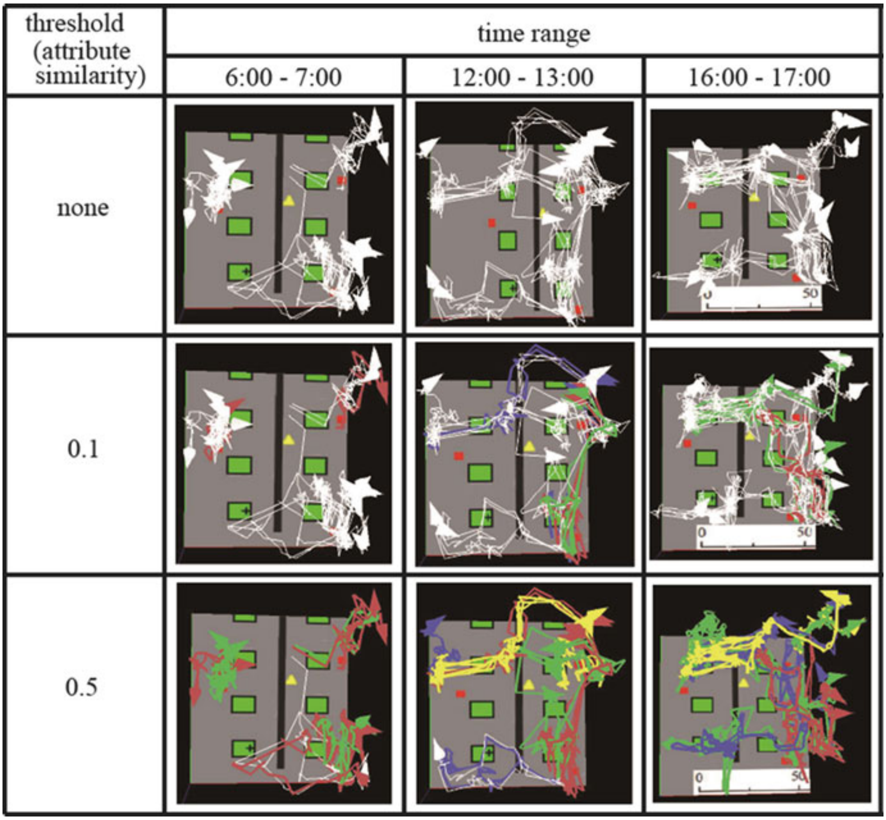


Fig. 21 Attribute similarity and time ranges

chickens that follow other chickens throughout a day. These are mainly hens that follow a cock that leads his own chicken group. The attribute similarity function may have detected this type of flocking behavior.

5 Discussion and Conclusion

This research develops an interactive visualization toolkit based on temporal granularity and spatial similarity to explore and discover multi-scale mobility patterns in mobile objects databases. The toolkit facilitates highly interactive visual exploration of mobile trajectories at varying levels of temporal granularity and thresholds for trajectory aggregation based on locational and attribute similarity among paths at the specified granularity level. A case study of wild chicken mobility dataset shows that combination of both time granularity and trajectory aggregation facilitates mobility pattern detection. The interactive temporal aggregation method with OLAP in the proposed toolkit is the first step to explore trajectory data to mitigate the difficulty of exploring complex movement patterns. In addition, visualization with similarity thresholds provides distinct views of the movement data, with some discovered patterns being robust across different granularities and others being dependent on these parameters. The flexibility of the temporal querying and trajectory aggregation also allows the discovery of temporally recurrent mobility patterns: both locational and attribute similarity measures detected mobility patterns that Okabe et al. (2006) also uncovered with the same dataset.

There are some remaining challenges for continued development of the mobility visualization toolkit. Although the similarity functions in this research are scalable and effective at detecting similar mobility patterns, there are many ways to assess trajectory similarity. Other ways of measuring locational and attribute similarity should be explored, as well as other definitions of path similarity distinct from the two dimensions explored in this research. One possible way is to examine the properties that can be extracted from trajectories. There are other characteristics that can be calculated from trajectories other than five characteristics described as attribute similarity such as average travel distance, average x coordinate location and y coordinate location of a trajectory, and so on. Behavioral characteristics such as number of activities, number of visiting locations activity duration time are also candidates.

The trajectory summarization methods in the toolkit consider only the central tendency (mean values) of trajectory parameters such as velocity and direction. Searching for patterns based on central tendencies is reasonable for exploratory analysis since these patterns should be tested using confirmatory techniques before being accepted as knowledge. Nevertheless, a more complete representation of trajectory similarity would consider the dispersion (variance) of these parameters. A research frontier is to incorporate parameter variance in the summarization methods in a manner that is both scalable and intuitive to the analyst.

A related research challenge is linking the scalable, exploratory tools in this research to confirmatory techniques. The patterns discovered using the visualization methods are only hypotheses: these should be tested using more powerful analytical and statistical methods. These tools could be used in conjunction with the techniques in this toolkit to confirm and further analyze the tentative patterns discovered through visual exploration.

References

Alt H, Knauer H, Wenk C (2003) Comparison of distance measures for planar curves. Algorithmica 38:45–58

Andrienko G, Andrienko N (2008) Spatio-temporal aggregation for visual analysis of movements. In: IEEE symposium on Visual Analytics Science and Technology (VAST), Columbus, 21–23 Oct 2008, pp 51–58

Andrienko N, Andrienko G, Pelekis N, Spaccapietra S (2008) Basic concepts of movement data. In: Giannotti F, Pedreschi D (eds) Mobility, data mining and privacy: geographic knowledge discovery. Springer, Berlin, pp 15–38

Andrienko G, Andrienko N, Rinzivillo S, Nanni M, Pedreschi D, Giannotti F (2009) Interactive visual clustering of large collections of trajectories. In: Proceedings of the IEEE symposium on Visual Analytics Science and Technology (VAST), IEEE Computer Society Press, pp 3–10

Brimicombe A, Li Y (2006) Mobile space-time envelopes for location-based services. Trans GIS 10:5–23

Dodge S, Weibel R, Lautenschütz AK (2008) Towards a taxonomy of movement patterns. Inform Vis 7:240–252

Eiter T, Mannila H (1994) Computing discrete Fréchet distance. Technical University of Wien, CD-TR 94/64

Ester M, Kriegel HP, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases. In: Proceedings of 2nd international conference on knowledge discovery and data mining, Portland, pp 226–231

Gray J, Chaudhuri S, Bosworth A, Layman A, Reichart D, Venkatrao M, Pellow F, Pirahesh H (1997) Data cube: a relational aggregation operator generalizing group-by, cross-tab and sub-totals. Data Mini Know Disc 1:29–53

Hagerstrand T (1970) What about people in regional science? Papers Regional Sci Assoc 24:7–21

Han J, Kamber M (2006) Data mining – concepts and techniques. Morgan Kaufmann, San Francisco

Han J, Kamber M, Tung AKH (2001) Spatial clustering methods in data mining. In: Miller HJ, Han J (eds) Geographic data mining and knowledge discovery, 2nd edn. Taylor and Francis, London, pp 74–109

Han J, Lee JG, Kamber M (2009) An overview of clustering methods in geographic data analysis. In: Miller HJ, Han J (eds) Geographic data mining and knowledge discovery, 2nd edn. Taylor and Francis, London, pp 149–187

Hornsby KS (2001) Temporal zooming. Trans GIS 5:255–272

Hornsby KS, Cole S (2007) Modeling moving geospatial objects from an event-based perspective. Trans GIS 11(4):555–573

Hornsby KS, Egenhofer M (2002) Modeling moving objects over multiple granularities, special issue on spatial and temporal granularity. Ann Math Artif Intel 36:177–194

Huang Y, Chen C, Dong P (2008) Modeling herds and their evolvements from trajectory data. In: Proceedings of the 5th international conference on geographic information science (GIScience 2008), Park City, pp 90–105

Huttenlocher DP, Klanderman GA, Rucklidge WJ (1993) Comparing images using the Hausdorff distance. IEEE Trans Pattern Anal Mach Intell 15(9):850–863

Kapler T, Wright W (2005) GeoTime information visualization. Inform Vis 4:136–146

Kraak M-J (2003) The space-time cube revisited from a geovisualization perspective. In: Proceedings of the 21th international cartographic conference, Durban, pp 1988–1995

Kraak MJ, Huisman O (2009) Beyond exploratory visualization of space-time paths. In: Miller HJ, Han J (eds) Geographic data mining and knowledge discovery. CRC, Boca Raton, pp 431–443

Kwan M-P (2000) Interactive geovisualization of activity-travel patterns using 3D geographical information systems: a methodological exploration with a large data set. Transport Res Part C 8:185–203

Langran G (1992) Time in geographic information systems. Taylor and Francis, London

Laube P, Imfeld S, Weibel R (2005) Discovering relative motion patterns in groups of moving point objects. Int J Geogr Inform Sci 19:639–668

Leonardi L, Marketos G, Frentzos E, Giatrakos N, Orlando S, Pelekis N, Raffaeta A, Roncato A, Silvestri C, Theodoridis Y (2010) T-warehouse: visual olap analysis on trajectory data. In: IEEE 26th international conference on data engineering (ICDE), 2010, Long Beach, pp 1141–1144. IEEE

Li C, Longley P (2006) A test environment for location-based services applications. Trans GIS 10:43–61

Lu C-T, Boedihardjo AP, Shekhar S (2009) Analysis of spatial data with map cubes: highway traffic data. In: Miller HJ, Han J (eds) Geographic data mining and knowledge discovery, 2nd edn. Taylor and Francis, London, pp 223–245

Miller HJ (1991) Modelling accessibility using space-time prism concepts within geographical information systems. Int J Geogr Inform Syst 5:287–301

Miller HJ (2005) A measurement theory for time geography. Geogr Anal 37:17–45

Miller HJ, Han J (2009) Geographic data mining and knowledge discovery: an overview. In: Miller HJ, Han J (eds) Geographic data mining and knowledge discovery, 2nd edn. Taylor and Francis, London, pp 3–32

Okabe A, Satoh T, Okabe K, Imamura E, Morathop S, Jailanka C, Ratanasermpong S, Hayashi Y, Akishinonomiya F (2006) Acquisition of spatio-temporal data of free-range chicken using a Wireless Fidelity (WiFi) positioning system and spatial analysis of the obtained data. The papers and proceedings of the geographic information systems association, vol 15, pp 395–400

Peuquet DJ, Duan N (1995) An event-based spatiotemporal data model (ESTDM) for temporal analysis of geographical data. Int J Geogr Inform Syst 9:7–24

Pfoser D, Theodoridis Y (2003) Generating semantics-based trajectories of moving objects. Comput Environ Urban Syst 27:243–263

Rinzivillo S, Pedreschi D, Nanni M, Giannotti F, Andrienko N, Andrienko G (2008) Visually driven analysis of movement data by progressive clustering. Inf Vis 7(3-4):225–239

Roberti M (2003) Wal-Mart spells out RFID vision. [online]. RFID Journal. http://www.rfidjournal.com/. Accessed 16 June 2003

Schreck T, Bernard J, Tekusova T (2008) Visual cluster analysis of trajectory data with interactive Kohonen Maps. In: Proceedings of the IEEE symposium on visual analysis science and technology (VAST 2008), Columbus, pp 3–10

Sedgewick R (1990) Finding the convex hull. In: Sedgewick R (ed) Algorithms in C. Addison-Wesley Publishing Company, Reading, pp 359–372

Shao F, Cai S, Gu J (2010) A modified Hausdorff distance based algorithm for 2- dimensional spatial trajectory matching. In: The 5th international conference on computer science & education, 24–27 Aug, Hefei, pp 166–172

Shaw S-L, Yu H, Bombom LS (2008) A space-time GIS approach to exploring large individual-based spatiotemporal datasets. Trans GIS 12:425–441

Shekhar S, Lu CT, Tan X, Chawla S, Vatsavai RR (2001) A visualization tool for spatial data warehouses. In: Miller HJ, Han J (eds) Geographic data mining and knowledge discovery. Taylor and Francis, London, pp 74–109

Shekhar S, Lu CT, Liu R, Zhou C (2002) CubeView: a system for traffic data visualization. Intelligent transportation systems. In: Proceedings of the fifth IEEE international conference on intelligent transportation systems, pp 674–678

Sinha G, Mark DM (2005) Measuring similarity between geospatial lifelines in studies of environmental health. J Geogr Syst 7:115–136

Sistla P, Wolfson O, Chamberlain S, Dao S (1998) Querying the uncertain position of moving objects. In: Etzion O, Jajodia S, Sripada S (eds) Temporal database: research and practice. Springer, New York, pp 310–337

Skupin A (2008) Visualizing human movement in attribute space. In: Agarwal P, Skupin A (eds) Self-organizing maps: applications ain geographic information science. Wiley, Chichester

Skupin A, Hagelman R (2005) Visualizing demographic trajectories with self-organizing maps. GeoInformatica 9(2):159–179

Steiner I, Bürgi C, Werffeli S, Dell'Omo G, Valenti P, Tröster G, Wolfer DP, Lipp H-P (2000) A GPS logger and software for analysis of homing in pigeons and small mammals. Physiol Behav 71:589–596

Turchin P (1998) Quantitative analysis of movement: measuring and modeling population redistribution in animals and plants. Sinauer Associates, Sunderland

Vlachos M, Kollios G, Gunopulos D (2002) Discovering similar multidimensional trajectories. In: Proceedings of the 18th international conference on data engineering, 2002, San Francisco, pp 673–684

Wentz EA (2000) A shape definition for geographic applications based on edge, elongation, and perforation. Geogr Anal 32–1:95–112

Wentz EA, Campell AF, Houston R (2003) A comparison of two methods to create tracks of moving objects: linear weighted distance and constrained random walk. Int J Geogr Inf Sci 17:623–645

Wood J, Dykes J (2008) Spatially ordered treemaps. IEEE Trans Vis Comput Graph (Proceedings Visualization/Information Visualization 2008), 14(6):1348–1355

Worboys M, Duckham M (2004) GIS: a computing perspective, 2nd edn. Taylor and Francis, London

Yanagisawa Y, Akahani JI, Satoh T (2003) Shape-based similarity query for trajectory of mobile objects. In: Chen M-S et al (eds) Mobile data management, LNCS 2574. Springer, Berlin, pp 63–77

Yu H, Shaw S-L (2008) Exploring potential human activities in physical and virtual spaces: a spatio-temporal GIS approach. Int J Geogr Inf Sci 22:409–430

Yuan M (2001) Representing geographic information to support queries about life and motion of socio-economic units. In: Frank A, Raper J, Cheylan J-P (eds) Life and motion of socio-economic units. Taylor and Francis, London, pp 217–234