12

Air Traffic Analysis

Christophe Hurter, Gennady Andrienko, Natalia Andrienko,
Ralf Hartmut Güting, Mahmoud Sakr

12.1 Introduction

The goal of Air Traffic Control (ATC) is to maximize both safety and capacity, so as to accept all flights without compromising the life of the passengers or creating delays. Because air traffic is expected to double by 2030, new visualizations and analysis tools have to be developed to maintain and further improve the safety level. To do so, air traffic practitioners analyze data from the ATC activity. This multidimensional data includes aircraft trajectories (3D location plus time), flight routes (ordered sequences of spatio-temporal points that represent planned routes), meteorological data, etc. In this chapter, we detail the relevant tasks of ATC practitioners, and demonstrate recent visualization and query methods to fulfill them.

The special properties of ATC data propose new challenges and, at the same time, new opportunities of data analysis. The semantics of the data is rich because it includes the third dimension (altitude), which can be used to discover salient events such as take offs and landings. More semantics can be added by augmenting background data such as the traffic network and the meteorological data. ATC data sets are characterized by their large sizes, adding more challenges to the analysis. Trajectory analysis is difficult due to the data set size and to the fact that it contains many errors and uncertainties. One day's traffic over France contains about 20000 trajectories (> 1 million records). Recording is done in a periodic manner (in our database: a radar plot, per aircraft, every 4 minutes), but a plot can be missed, or have erroneous data because of physical problems that occur at the time of recording.

This chapter demonstrates recent works of trajectory analysis. Three techniques are demonstrated: direct manipulation, visual analytics, and moving object database queries. Direct manipulation visually represents the raw trajectories, and allows the user to efficiently explore them and highlight interesting subsets using convenient views and simple mouse interaction. Visual analytics provide a rich tool box of data transformations and visualizations that help a human analyst exploring complex movement events in the data. Moving object database (MOD) defines query operators accessible to the user through textual query languages. They are able to perform complex computations over large data sets efficiently. According to the analysis task, the experience of the human doing the analysis, and the data set size, any of these three analysis methods (or a combination of them) can be chosen. This is illustrated in Figure 12.1.

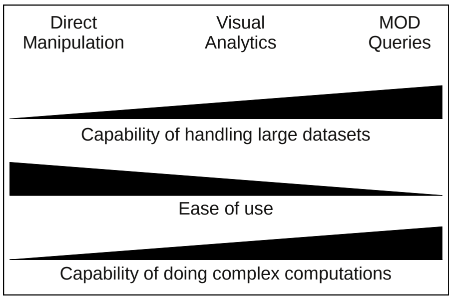


Figure 12.1 Factors of choosing among the methods of trajectory analysis.

Direct manipulation is good for having a first look at the data. It is intuitive to use. Visual analytics provides more sophisticated transformations and aggregations, and thus it is able to process larger data sets, and to perform deeper analysis. Human expertise is, however, a deciding factor for good analysis results. MOD queries are mandatory for complex computations, such as pattern matching. The user must however know exactly what he or she is looking for, and how to precisely describe it in terms of the MOD query language.

Throughout this chapter we will demonstrate each of these analysis methods, in the context of real tasks, and using a real data set. The motivation for the analysis, and the description of the data set are presented in Sections 12.2 and 12.3 respectively. Direct manipulation is demonstrated in Section 12.4. Section 12.5 demonstrates the use of visual analytics to explore movement events, such as landings and take-offs, and to derive useful statistics from them. Finally Section 12.6 explains a MOD query operator that is able to match complex patterns in ATC data, such as missed approaches and step-wise descents.

12.2 Motivation

Aircraft trajectories are monitored and recorded by ground radar. It is displayed in real-time on radar screens. This data is essential for *air traffic controllers*, in order to maintain a safe distance between aircraft and to optimize traffic fluidity (reduce flight time, noise and fuel consumption). Our goal in this chapter is not to provide tools for real-time usages, but rather to detail off-line tools that analyze recorded trajectories in more depth. Without this real-time constraint, ATC practitioners can investigate, in more detail, recorded trajectories and therefore extract relevant information and perform three main tasks: improve safety, optimize traffic, and monitor environmental considerations.

Improving safety can be detailed as:

1 analyze and understand past conflicts (when two aircraft fail to meet minimum safety distance) and then improve safety with feedback from past experience,
2 analyze the accuracy of data provided by ground radar with probe trajectory comparison (i.e., with GPS tracking and radar test plots), and
3 filter and extract trajectories in order to reuse them for Air Traffic Controllers' training simulations.

Traffic optimization can be detailed as:

1 devise new air space organization and flight routes to handle traffic increase,
2 study profitability (i.e., number of aircraft on a specific flight route per day, number of aircraft that actually land at a specific airport, etc),
3 calculate the metrics from the traffic: traffic density, spacing quality (mean distance between aircraft), number of holding loops, number of rectilinear trajectories (trajectories that are close to the shortest path departure - arrival), etc, and
4 measure the activity of each airport: number of take-offs and landings per hour etc.

Finally environmental considerations can be detailed as:

1 Compare trajectories with environmental considerations (fuel consumption, noise pollution, vertical profile comparison),
2 detect missed approach trajectories (which produce noise), lap training landings (pilots who train to take off, fly around the air field and land. Lap training landings consume a lot of fuel), and
3 count continuous descending aircraft (since these aircraft maintain a constant descent rate, they reduce their fuel consumption).

This list is not exhaustive but it gives the main tasks that ATC practitioners perform. These tasks highlight the need for powerful tools to analyze aircraft trajectories.

12.3 Data Set Description

In this section, we detail the different steps required to produce data sets of aircraft trajectories provided by the IMAGE system. In France, ground radars send aircraft positions through the RENAR (RÉseau de la Navigation AéRienne) network. Due to network bandwidth limitations we cannot route all raw radar information toward a single network access point to record it. Therefore, we use the French IMAGE system. IMAGE is a system that aims to gather aircraft positions from all French controlled areas. Its goal is neither to monitor aircraft activity, nor to optimize traffic flow, but to give a general view of the traffic (communication purposes). The IMAGE system is connected to the five French STRs (Système de Traitement Radar), one in each en-route control center (Figure 12.2). STR systems receive aircraft information from different radar sources and calculate an estimated position for each monitored aircraft (using tracking and smoothing algorithms). The IMAGE system helps to reduce ground radar sources to only 5 data sources, and enables us to retrieve aircraft positions over France within the RENAR network.

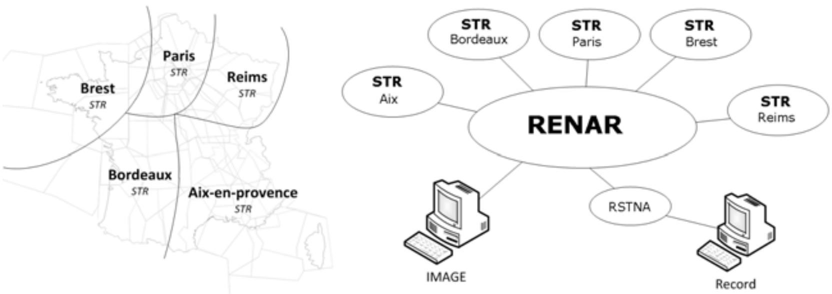


Figure 12.2 IMAGE network with STRs.

Merging the 5 data sources raises lots of issues: unique aircraft identifiers, overlapping areas, time stamps and sampling rates. Firstly, each STR sends the aircraft position with an identifier from 1 to 1023. Since more than 1023 aircraft can fly over France at the same time, we extend this identifier to a 16-bit format and re-reassign a unique identifier to every trajectory. To do so, we use a spatio-temporal frame filtering to assign a new unique identifier to each trajectory: each radar plot that has the same identifier within a 600 second time frame within an area of 200 km (100 Nm, Nautical Miles) radius (which corresponds to a 12 minute straight flight at high altitude) belongs to the same trajectory. At this stage, trajectories with less than three plots are removed and no trajectory has the same identifier.

Secondly, we merge all the 5 new, re-assigned, radar records into one file. The main issue is to connect trajectories that were recorded by different STR sources. To do so, we resample all the data to insure that every record has the same regular time-stamp. Then we setup the following merging parameters: when two trajectories overlap, they merge if the overlapping points have the same altitude (less than 600 m/2000 ft, which corresponds to 1 minute descent), and close position (less than 9 km/5 Nm, which corresponds to the minimal safety distance).

The properties of the data set we use in this chapter are typical to any IMAGE data set. We use a data set with 17,851 flight trajectories over France during one day (Friday, the 22nd of February, 2008) consisting of 427,651 records. The trajectories, shown in Figure 12.3, include flights of passenger, cargo, and private airplanes and helicopters. The temporal resolution of the data mostly varies from 1 to 3 minutes, although larger time gaps (up to 5 minutes) also occur. 3,000 trajectories (60,000 records, 16%) fly over France per day without landing.

12.4 Direct Manipulation of Trajectories

Formulating trajectory queries are difficult for two reasons. Firstly, they are often only specifiable with visual features (straight lines, or general shapes). Secondly, users often explore the queries as much as they explore the data: in the course of exploration, users discover that the set of features they thought relevant has to be adapted, either because they were false, or because they cannot find how to query them efficiently. Furthermore, trajectories are numerous and tangled: one-day's traffic over France for example, represents some 20000 trajectories. When dealing with trajectories, users must perform dynamic requests (response time < 100 ms) on a large multi-dimensional data set (>1 million data) which contains many errors and uncertainties. The problem we address in this section is to find a way to express these queries, simply and accurately, given the constraints of size and uncertainty of the data sets. As a solution, the visualization and direct manipulation of trajectories proposes efficient interactions features. Direct manipulation was introduced by Ben Shneiderman in 1983 (Shneiderman (1983)) within the context of office applications and the virtual desktop metaphor. This term has been extended to human-computer interaction paradigms. The intention is to allow users to directly manipulate objects presented to them, using actions that correspond to the physical world (e.g., grasp, move objects, etc).

In the following sections, we first describe direct manipulation requirements for trajectory exploration, then we detail an implementation instance, and finally we give one scenario of usage.

12.4.1 Design Requirements for Trajectory Exploration

Based on trajectory data set characteristics, we extracted the following design requirements to achieve the visual exploration of trajectories:

1 View configuration: the system must permit the customization of views so as to offer multiple means of understanding and visually querying the data. It should allow for a change of mapping between data and visual dimensions. The system should also provide smooth transitions between visual configurations. Hence, the user will be able to visually track patterns between different view configurations.

2 Views organization and navigation: the system must also permit the display of multiple views. The user must be able to visually compare different visual configurations of the data set. This can be done with a matrix scatterplot or juxtaposed views.

3 View filtering: the system must allow the user to filter out trajectories and then reduce cluttering.

4 Trajectory selections and Boolean operations: The system must enable the user to select trajectories and combine them in order to perform complex queries. Some systems allow multiple selections sometimes called “layers”. Users can combine layers with Boolean operation by applying an “and” operation when they try to group differently selected trajectories.

12.4.2 Implementation Instance: FromDaDy

We have developed *FromDaDy* (Hurter et al. (2009)) (which stands for “FROM DAta to DisplaY”), a visualization tool that tackles the challenge of representing, and interacting with numerous trajectories (several million trajectories composed of up to 10 million points). FromDaDy employs a simple paradigm to explore multidimensional data based on scatterplots, brushing, “pick and drop”, juxtaposed views and rapid visual configuration. Together with a finely tuned mix between design customization and simple interaction, users can filter, remove and add trajectories in an incremental manner until they extract a set of relevant data, thus formulating complex queries.

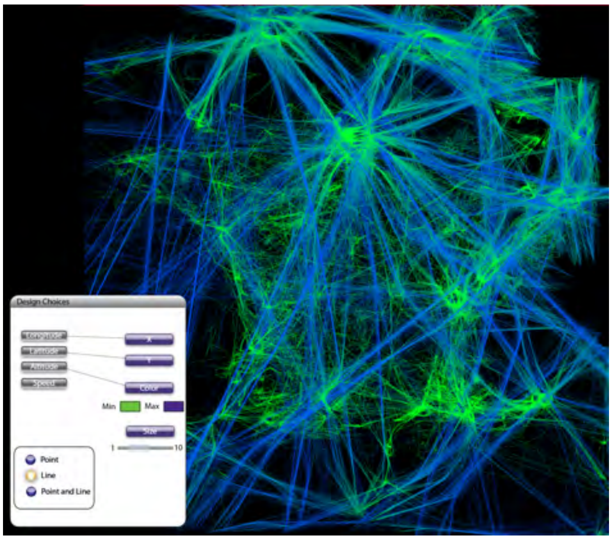


Figure 12.3 One day's record of traffic over France. The color gradient from green to blue represents the ascending altitude of aircraft (green being the lowest and blue the highest altitude). The French coastline is apparent here in terms of pleasure flights by light aircraft and the straight blue lines represent high altitude flight routes. A user interface shows the data set fields and the defined visual configuration.

12.4.3 Views Organization and Navigation

A FromDaDy session starts with a view displaying all the data in one scatterplot. The visualization employs a default visual configuration, e.g., the mapping between data dimensions and visual variables. The view is inside a window, and occupies a cell in a virtual infinite grid that extends from the four sides of the cell. The user can configure the two axes of each scatterplot and use other visual variables such as color and line width to display data set dimensions. For instance, in Figure 12.3, the user attached the data set field latitude to the Y axis, and the field longitude to the X axis. The user also chose to use the altitude to color trajectory sections, such as, low altitudes in green and high altitudes in blue.

12.4.4 Trajectory Manipulation

We have implemented a simple and efficient direct manipulation technique: trajectory brush, pick and drop. The user selects a subset of the data set by means of a brushing technique. Brushing is an interaction that allows the user to "brush" graphical entities, using a size-configurable or shape-configurable area controlled by the mouse pointer. Each trajectory touched by this area is selected, and becomes gray. The selection can be modified by further brush strokes, or by removing parts of it with brush strokes in the "erase" mode. The display shows a brush trail, so that the user can see and remember more easily how the selection was made. The combination of fast switching between the add/erase mode, trajectory visualization, rapid size-setting, and cursor-centered zooming allows for fast and incremental selection.

Then the user can pick bushed trajectories by hitting the space bar. The user extracts previously selected data from the current scatterplot and attaches it to the mouse pointer so it appears in a "fly-over" view (transparent background). When the user hits the space bar for the second time, a drop occurs in the view under the cursor. If the view under the mouse pointer is empty, the software creates a new scatterplot with the selected data. If the user presses the space bar while moving over a view containing data, FromDaDy adds the selected data to this scatterplot. Although it resembles a regular drag and drop operation, we prefer to use the term "pick and drop", because the data is removed from the previous view and is attached to the cursor even if the space bar is released. The user can also destroy a view if the brush selects all the trajectories and the user picks them.

12.4.5 Brush Pick and Drop

The fundamentally new aspect of FromDaDy compared to existing visualization systems, is to enable users to spread data across views. Within FromDaDy, there is a single line displayed per trajectory: trajectories are not duplicated, but are spread across views. The advantage of this technique is multifold. It enables the user to remove data from a view (and drop it on to the destination view). The fly-over view enables the user to rapidly decide if the revealed data (previously hidden by the picked data) is interesting. Secondly, it makes it possible to build a data subset incrementally. In this case, the user can immediately assess the quality of the selection, by seeing it in the "fly-over" view. Furthermore, by removing data from the first view, the user makes it less cluttered, and this makes it easier for them to pick and drop more trajectories.

Another advantage of the brush pick and drop paradigm is that this interaction helps the user to perform complex Boolean operations: "I want the trajectories that go into this area but not the ones that are too high and only those that are faster than a given minimum speed." A seminal previous work uses containers (also called layers) to cluster trajectories and explicitly applies Boolean operations to combine them. Even with an astute interface, Boolean operations are cumbersome to produce, since results are difficult to foresee. FromDady overcomes this drawback, since all the operations of the interaction paradigm (brush, pick and drop) implicitly perform Boolean operations. Removing trajectories corresponds to an XOR operation and dropping trajectories corresponds to an ADD operation. The following examples illustrate the union (AND), intersection (OR) and negation (NOT) Boolean operations. With these three basic operations the user can perform all kinds of Boolean operations: AND, OR, NOT, XOR, etc.

In Figure 12.4, users want to select trajectories that pass through region A or through region B. They just have to brush the two desired regions and Pick/Drop the selected tracks into a new view. The resulting view contains their query, and the previous view contains the negation of the query.

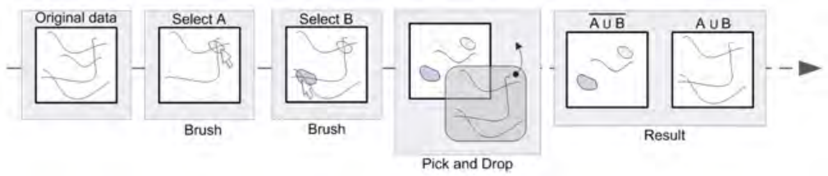


Figure 12.4 Union Boolean operation.

12.4.6 Example of Usage

In this scenario, we use one day recordings of aircraft trajectories over France. In this data set, a unique and incremental identifier is assigned to each trajectory. The first trajectory of the day has the number 0', the next one has the number 1' etc. Figure 12.5 shows an abstract visualization of this data set. The X screen axis shows the time of each radar plot and the Y screen axis shows the aircraft's identifier. Since these identifiers are incremental over the day, the resulting visualization shows a noticeable continuous shape, in which each horizontal line represents the duration of one flight. The slope of the shape indicates the traffic increase during the day (due to the incrementally assigned identifiers). Hence, the traffic notably increases at 5 am and decreases at 10 pm as reflected in the change of slope. The width of this shape indicates the

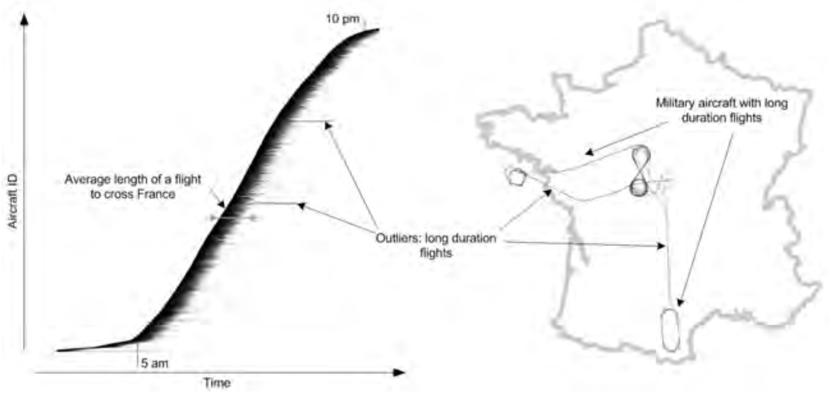


Figure 12.5 Detection of supply planes with an abstract visualization.

average flight duration in the data set: it is about 2.5 hours which represents the average time taken to cross France. But some aircraft have longer trajectory durations. The user brushes these long trails (the ones that come out of the curved shape). When visualizing them with a latitude (Y screen) and longitude (X screen) visual configuration, the user discovers a figure of eight shaped trajectory. This trajectory covers 6 hours and performs 11 loops. After further investigation, it is found that it corresponds to a military supply plane.

This data exploration has been done with a visualization tool. The user would have also been able to perform the same extraction with a textual tool, like SQL queries. The only difference is that, a textual tool would not have led the user to the idea of exploring long flight duration in order to extract military aircraft. Only with the incremental trajectory exploration, can the user discover the valid requests for this data set. In a sense, the user explores the data set, and at the same time, explores the request to perform. Even if this process is efficient, the direct manipulation cannot be automatic. Analysts need tools to enhance their exploration capabilities. Therefore, extended work will be presented in the following sections.

12.5 Event Extraction

There is a class of problems where analysts need to determine places in which movement events (m-events) of a certain type repeatedly occur and then use these places in further analysis. The relevant places can only be delineated by processing movement data, that is, there is no predefined set of places (e.g., compartments of a territory division) from which the analyst can select places of interest. The relevant places may have arbitrary shapes and sizes and irregular spatial distribution. They may even overlap in space; therefore, approaches based on dividing the territory into non-overlapping areas, as in Andrienko and Andrienko (2011), are not appropriate. In this section, we analyze one-day record of aircraft trajectory with a visual analytics procedure for place-centered analysis of mobility data (Andrienko et al. (2011c)). The procedure consists of four steps: (1) visually-supported extraction of relevant m-events, (2) finding and delineating significant places on the basis of interactive clustering of the m-events according to different attributes, (3) spatio-temporal aggregation of the m-events and movement data by the defined places or pairs of places and time intervals; (4) analysis of the aggregated data for studying the spatio-temporal patterns of event occurrences and/or connections between the places.

12.5.1 Analyzing Flights Dynamics in France

We shall apply our visual analytics procedure to ATC data with the following goals: (1) Identify the airports in use. (2) Investigate the temporal dynamics of the flights to and from the airports (i.e., landings and takeoffs). (3) Investigate the connections among the airports, the intensity of the flights between them, and their distribution over a day.

It may not be obvious to the reader why the airport areas need to be determined from the data instead of using the official airport boundaries, which should be known. The problem is the low temporal resolution of the data. For many flights, the first recorded positions lie outside the boundaries of the origin airports and/or the last recorded positions are not within the boundaries of the destination airports. Therefore, to refer the flights to their origin and destination airports, it is necessary to build sufficiently large areas around the airports that would include the available first and last points. It is not known in advance how large the areas need to be and what geometrical shapes are appropriate.

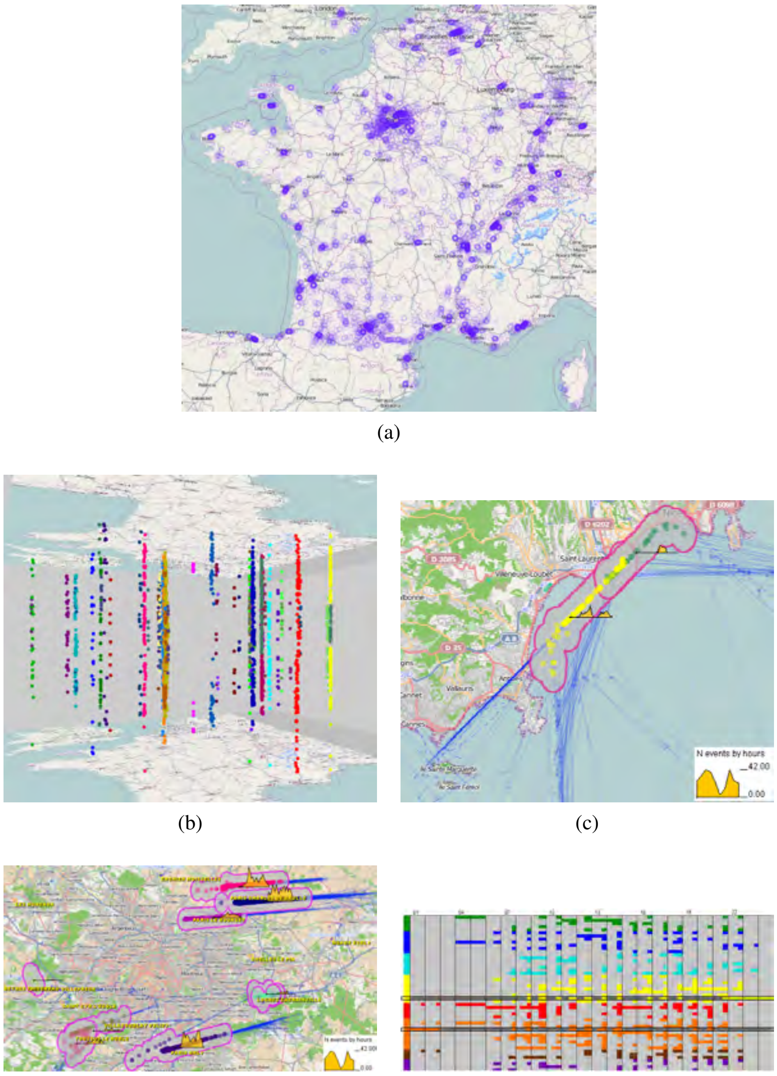
Our approach to defining the areas is based on the background knowledge that airplanes typically land and take off in similar directions, which are determined by the orientation of the airport runways. We extract the available last positions of the aircraft that landed and first positions of those that took off and cluster them by spatial positions and movement directions using a density-based clustering method Optics (Ankerst et al. (1999)) with similarity measures designed for spatio-temporal events (Andrienko et al. (2011c)). As a result, points lying outside or even quite far from the airports are grouped together with the points lying within the airport boundaries if they correspond to landings or takeoffs with similar directions. The airport "catchment" areas are built as buffers around these clusters. The areas can be verified using the known positions of the airports: they must be within the areas.

Not always do starts and ends of trajectories correspond to takeoffs and landings. The radar observation data also contain parts of transit trajectories that just pass over France as well as flights going outside France and those coming to France from abroad. Real takeoffs and landings must be distilled from the available starts and ends of the recorded tracks. To extract the landings, we use the following query condition: the altitude is less than 1 km in the last 5 minutes of the trajectory. From each trajectory that has such points, we extract the last point as an m-event representing the landing (Figure 12.6a). In the second step of the analysis, we cluster the landing events by the spatial positions and directions (SD) using the thresholds of 1 km and 30 degrees, respectively. The resulting SD-clusters are presented in the space-time cube in Figure 12.6b; the noise (events not having sufficient counts of SD-neighbors) is excluded. The colors represent different clusters. The vertical alignments of points correspond to the airports where multiple landings took place during the day.

An interesting pattern can be observed in the area of Nice on the Southeast of France. There are two SD-clusters of landings, yellow and green; their points make a column on the right in the cube. The green cluster appears as an intrusion inside the yellow one. This means that the landing direction changed in this area twice during the day due to a change of wind direction (aircraft take off and land facing the wind). The map fragment in Figure 12.6c shows that the yellow cluster contains landings from the Southwest and the green cluster landings from the Northeast. The blue lines in Figure 12.6 show the last ten minutes fragments of the respective trajectories and reflect the mandatory landing directions.

The observation of the direction changes gives us an idea that the temporal patterns of landings should be investigated not by airports only but by airports and landing directions. Therefore, we build 500-meter spatial buffers around the SD-clusters, as shown in Figure 12.6c. For an analysis by airports, irrespective of the directions, we would do a second stage of clustering (after excluding the noise) by only the spatial positions of the events and then build buffers around the resulting spatial clusters.

In the third step of the analysis procedure, we aggregate the landing m-events in space by the buffers and in time by one-hour intervals. In the fourth step, we visualize the resulting time series by temporal diagrams positioned on the map display; two of them can be seen in the map fragment in Figure 12.6c. They show that the aircraft landed in the airport of Nice from the Southwest almost all time except for an interval in the middle of the day, when the landing



(a)

(b) (c)

(d) (e)

Figure 12.6 Event extraction results. (a): The positions of the landing events extracted from the flight data are drawn with 50% opacity. (b): The space-time cube shows the landing events clustered by spatial positions and directions. (c): The yellow and green dots represent two SD-clusters of landings in the airport of Nice. The time diagrams show the dynamics of the landings from two directions. (d): The time diagrams show the dynamics of landings in the airports of Paris. (e): The flights distribution between the airports by hourly intervals. Highlighted are rows for the connections Marseille-Paris (yellow) and Paris-Marseille (orange).

direction changed to the opposite. The exact times and values are displayed when the mouse cursor points on an area.

Figure 12.6d presents the map with the temporal diagrams for the Paris region. We can see that the Orly airport and the Northern runway of the Charles de Gaulle airport have clear peaks in the morning and in the evening. It is a typical pattern for airline hubs: short period of time, during which many flights arrive and take off, maximize the number of possible connections. The Southern runway of the Charles de Gaulle airport is used with almost constant intensity during the day. The remaining airports are used much less intensively and mostly in the afternoon.

So far we have considered only the landings. To investigate the takeoffs, we repeat the procedure. To extract the takeoff events in the first step, we use the query condition that the altitude must be less than 1 km at the beginning of the trajectory. The remainder of the procedure is similar to that for the landings.

To investigate the connections among the airports, we need to define the airport areas so that they include both the takeoff and the landing events. We join the sets of the takeoff and landing events, which have been previously filtered by removing the noise after the SD-clustering. Then we apply clustering by spatial positions, to unite the clusters of takeoffs and landings in different directions occurring at the same airports. We build spatial buffers around the spatial clusters to obtain the airport areas. In the third step (spatio-temporal aggregation), we aggregate the trajectories by pairs of places (airport areas) and time intervals (1 hour length). We use only those trajectories that have both takeoff and landing events. As a result, we obtain aggregate flows (vectors) with respective hourly time series and totals of flight counts.

To investigate the aggregates (Step 4: analysis of the aggregated data), we visualize the total counts on a flow map. The aggregate flows are shown by directed arrows with the widths proportional to the flight counts. By interactive filtering, we hide minor flows (less than 5 flights) and focus on the short-distance flows (below 100 km distance). We see that there are quite many flights connecting close airports, particularly, in Paris. As explained by a domain expert, a part of them are flights without passengers used for relocating aircraft between big airports, such as Charles de Gaulle and Orly. Short-distance flows between small airports correspond to training and leisure flights of private pilots. Focusing on the long-distance flows (100 km and more) reveals a mostly radial connectivity scheme with a center in Paris.

To investigate the temporal dynamics of the flows, we use the table display as shown in Figure 12.6e. The columns of the table correspond to the hourly time intervals and the rows to the flows. The lengths of the colored bar segments in the cells are proportional to the flight counts for the respective flows and intervals. The colors correspond to the eight compass directions. The table view is linked to the flow map. Thus, clicking on the vectors connecting Paris Orly and Marseille on the map, we get two rows highlighted. The yellow one corresponds to the Northwestern direction, i.e., from Marseille to Paris, and the orange one to the opposite direction from Paris to Marseille. There are one or two flights from Marseille to Paris every hour in the intervals 07-14h and 15-18h and three flights per hour from 22h to midnight. The traffic in the opposite direction has a different profile: 3 flights per hour from midnight till 02h and several flights in the morning, at noon, and in the evening. The complementary link from the table view to the map can be used to locate flows with particular dynamics.

12.5.2 Validation of the Findings

First, to assess the validity of the extracted areas of takeoffs and landings, we compared them with the known positions of the airports and found that the areas include the airports. Furthermore, the areas have elongated shapes (Figure 12.6d) whose spatial orientations coincide with the orientations of the runways of the respective airports. Next, the results of data aggregation by the areas (i.e., counts of takeoffs, landings, and flights between airports) correspond very well to the common knowledge about the sizes and connectivity of the French cities and airports. The discovered patterns have been also checked and interpreted by a domain expert who confirmed their plausibility.

12.6 Complex Pattern Extraction Using a Moving Object Database System

Moving object database systems are another good candidate for air traffic analysis. This section demonstrates a concrete example of using the Secondo MOD system in order to extract complex spatio-temporal patterns from the flight trajectories. The task is to extract the *missed approach* and the *step-wise descent* events that occurred in the ATC data set described in Section 12.3. The *spatio-temporal pattern* (*STP*) *algebra* in Secondo brings a generic set of query operations accessible through the Secondo query languages to let the user express arbitrarily complex patterns and efficiently match them on large moving objects databases. This algebra defines the STP predicate, which is the main tool we are going to illustrate in the section. To get the most out of this section, please first read the chapter about moving object database systems (Chapter 3), especially the part explaining the Secondo query languages.

12.6.1 The Spatio-Temporal Pattern Predicate

A traditional select-from-where query is formulated based on a single predicate given in the where-clause. Such a query scheme is not sufficient when dealing with moving objects. A moving object has a life-time and it fulfills several predicates during it. In many applications it is required to find the objects that fulfill a set of predicates in a certain temporal order. In ATC, for instance, it is required to detect landing procedures such as *go-around*, *missed approach*, *touch-and-go*, etc. Each of these procedures consists of a set of well defined steps that have to be implemented by the pilot in a certain temporal order. Extracting these situations from the aircraft trajectories requires a query tool that accepts such descriptions, and matches them against the trajectories. Here comes the *spatio-temporal pattern predicate* to extend the traditional select-from-where scheme, and let the user formulate such queries.

Essentially the STP predicate is a pair $\langle P, C \rangle$, where P is a set of predicates and C is a set of temporal order constraints on their fulfillment. Given a tuple u, e.g., representing one flight trajectory, the STP predicate yields true iff u fulfills all the predicates in P in the temporal order asserted by all the constraints in C. Consider for example the *missed approach* procedure. It can be described by three predicates: aircraft comes close to destination, aircraft descends till a height of less than 1000 m, and aircraft climbs. Temporally, the third predicate must be fulfilled after the second predicate, and both of them must be fulfilled during the fulfillment time of the first predicate. Let's have a quick illustration of how this *missed approach* query is expressed using the Secondo executable language:

```
... stpattern[
  Close: distance(.Position, .Destination) < 5000.0,
  Down: ((.AltitudeDerivative < 0.0) and (.Altitude < 1000.0)),
  Up: .AltitudeDerivative > 0.0;
  stconstraint("Close", "Down", vec("abba","a.bba","baba")),
  stconstraint("Close", "Up", vec("abba","aba.b","abab")),
  stconstraint("Down", "Up", vec("aabb","aa.bb"))] ...
```

where `stpattern` is the Secondo operator denoting the STP predicate. For simplicity, we omit the query parts before and after the `stpattern` operator and denote them by three dots. The `stpattern` predicate is placed in the query as a filter condition within the Secondo `filter` operator. Here it receives a tuple with the schema:

tuple[`Id`: *int*, `Position`: *mpoint*, `Altitude`: *mreal*, `Destination`: *point*, `AltitudeDerivative`: *mreal*],

where `Position` represents the (Lon, Lat) of the aircraft and the `Altitude` is separately represented. This is because Secondo does not contain types for 3D moving points. The `Destination` is precomputed as the final (Lon, Lat) of the trajectory, and `AltitudeDerivative` is precomputed as the derivative of `Altitude`. The three predicates constituting P have the aliases `Close`, `Down`, and `Up`. The `Close` predicate asserts that the aircraft is close (within 5 km) to its destination airport. Note that this is a *time-dependent predicate*, also called *lifted predicate*. That is, the result of such a predicate is a time-dependent boolean *mbool*. It is false whenever the aircraft is far from its destination, and true whenever the aircraft is close to destination. Similarly, `Down` and `Up` are time-dependent predicates. Actually this is how the `stpattern` operator is able to check the temporal constraints on the predicate fulfillment, since an *mbool* contains information about when the predicate was fulfilled. The STP predicate expects that P be a set of time-dependent predicates, each of which is a mapping *tuple* → *mbool*. The aliases of the time-dependent predicates make it possible to refer to them in the temporal constraints.

The set of temporal constraints C in this example consists of the three temporal constraints denoted as `stconstraint`. Each of them asserts a temporal relation between two predicates forming a pair in P. The temporal relation is expressed by the `vec` operator. Each of the terms inside the `vec` operator specifies a relation between two time intervals. The start and the end points of the first interval are denoted `aa`, and those of the second interval are denoted `bb`. The order of the symbols describes the temporal order of the four end points. The dot symbol denotes the equality. For example, the relation `aa.bb` between the intervals i_1, i_2 denotes the order: $((i_1.t_1 < i_1.t_2) \wedge (i_1.t_2 = i_2.t_1) \wedge (i_2.t_1 < i_2.t_2))$. The temporal relation expressed by the `vec` operator is the disjunction of its components. A temporal constraint between two predicates p_i, p_j is fulfilled iff there exists an interval on which p_i is fulfilled, and another interval on which p_j is fulfilled, and the two intervals fulfill any of the interval relations in the constraint. For the STP predicate to be fulfilled, all the temporal constraints in C must be fulfilled.

Formally, given $P = \{p_1, ..., p_m\}$ a set of time-dependent predicates, $C = \{c_1, ..., c_n\}$ a set of constraints, and a tuple u, let $p_i(u)$ denote the evaluation of p_i for the tuple u (i.e., $p_i(u)$ is of type *mbool*). Let $[p_i(u)]_j$ denote the j^{th} time interval on which $p_i(u)$ is true. The evaluation of the STP predicate $\langle P, C\rangle$ for the tuple u is true iff: $\exists j_1..j_m$ such that the set of time intervals $[p_1(u)]_{j_1}..[p_m(u)]_{j_m}$ fulfills all the temporal constraints $c \in C$, and we call $[p_1(u)]_{j_1}..[p_m(u)]_{j_m}$ a *supported assignment*. The STP predicate yields true iff at least one supported assignment is found. This completes our description of the STP predicate.

The *STP Algebra* in Secondo defines other variants of the STP predicate (e.g., `stpatternexextendstream`). This operator is a triple $\langle P, C, f\rangle$ where

P, C are the same as before, and f is an additional condition on the time intervals of the supported assignments. One can express, for instance, that the `Down` predicate in this query must be fulfilled for at least 2 minutes. The `stpatternexextendstream` is also a stream operator, not a predicate. It extends every input tuple with attributes containing the time intervals on which the pattern occurs. Since one trajectory might contain several matches of the pattern, the `stpatternextendstream` copies the tuple, and extends every copy with one match. The following example expresses the *step-wise descent* scenario:

```
1  ...stpatternexextendstream[
2      Dive1: .SecondAltitudeDerivative < 0.0,
3      Lift: .SecondAltitudeDerivative >= 0.0,
4      Dive2: .SecondAltitudeDerivative < 0.0  ;
5      stconstraint("Dive1", "Lift", vec("aa.bb")),
6      stconstraint("Lift", "Dive2", vec("aa.bb"));
7      (end("Lift") - start("Lift")) > OneMinute ]
8    filter[isdefined(.Dive1) and
9      (AverageDiveAngle(.Alt atperiods .Lift) < 30.0)]...
```

In this scenario, the aircraft alternates between dive and cruise during its final approach. It is expressed as a sequence of increasing, decreasing, then again increasing rate of descent. Line 7 asserts that the `Lift` event stays more than a minute. Line 9 invokes the Secondo function object `AverageDiveAngle` to assert that the aircraft is flying almost horizontally during the `Lift` event, having a slope of less than 30° with the horizontal. The two queries in this section finish in approximately one minute on the given data set with 17,851 trajectories (427,651 records). The Secondo relation storing these flight trajectories occupies approximately 172 MB of disk-space on a Linux 32 bit machine.

12.6.2 Exploring Patterns by Integrating MOD with Visual Analytics

So far, we have shown that the STP predicates and its variants are very flexible and can be used to express arbitrarily complex patterns. In practice, tuning the parameters of these operators is tricky. The integration with visual analytics allows for fine tuning these parameters through user interaction. Secondo and V-Analytics realize such an integration scheme. They are integrated, so that it be possible to interchange query results in both directions. Typically the user starts by loading the whole data set in the databases of the two systems. The exploration starts in V-Analytics by removing incomplete data and artifacts, and sending the identifiers of the candidate trajectories to Secondo. In Secondo the user issues an STP query, and moves the result back to V-Analytics for validation. The visualization in V-Analytics helps the human analyst refining the query parameters. It can take as many cycles as needed between SECONDO and V-Analytics till the results are satisfactory.

The STP query can be written in SECONDO so that the result contains the time intervals in which the pattern occurred. These can be interpreted as movement events (m-events) in V-Analytics, so that the analysis procedures in the previous section are applicable. For example, one is able to explore the percentage of step-wise descents during one day, the percentage of missed approaches for each airport, the temporal distribution of missed approaches for a given airport, etc.

12.7 Conclusions

In this chapter, we gave an overview of up-to-date research techniques to explore and analyze trajectories. We detailed our motivations, gave the process we used to build trajectory data set, and explained three trajectory exploration techniques (direct manipulation, m-event, and MOD queries).

First, we introduced FromDaDy, a multidimensional visualization tool making it possible to explore large sets of aircraft trajectories with direct manipulation techniques. It uses a minimalist interface: a desktop with a matrix of cells, and a dimension-to-visual variables connection tool. Its interactions are also minimalist: brushing, picking, and dropping. Nevertheless the combination of these interactions permits numerous functions: the creation and destruction of working views, the initiation and refinement of selections, the filtering of data sets, the application of Boolean operations. The cornerstone of FromDaDy is the trajectory spreading across views with a simple brush/pick/drop paradigm. With the incremental trajectory exploration and direct manipulation, can the user discover the worthwhile requests for data sets. In a sense, the user explores the data set, and at the same time, explores the request to perform.

Second, we detailed a generic procedure for analyzing mobility data that is oriented to a class of problems where relevant places need to be determined from the mobility data in order to study place-related patterns of events and movements. The procedure includes: (1) extraction of relevant events from trajectories by queries involving diverse instant, interval, and cumulative characteristics of the movement and relations between the moving objects and elements of the spatio-temporal context, (2) density based clustering of the events by spatial positions, temporal positions, movement directions and, possibly, other attributes, which may be done in two stages for an effective removal of noise and getting clear clusters, (3) spatio-temporal aggregation of events and trajectories by the extracted places; and (4) analysis of the aggregated data. Visual analytics and m-events provide a rich tool box of data transformations and visualizations which help a human analyst exploring the data.

Third, MOD queries deal efficiently with vary large data sets with theoretically no limitation, and are able to express complex queries (neighborhood, patterns, aggregations, etc). While direct manipulation is easy to use (users are accustomed to manipulate tangible objects), it does not support automatic exploration. Furthermore, direct manipulation techniques need to be interactive which works again the data size. For instance FromDaDy can display up to 10 million points with an acceptable frame rate. If more data need to be displayed or manipulate, new computation technique need to be developed.

Since our visual analytics process uses m-events (geographic and temporal event), this tool is not suitable for complex computations like pattern extraction. MOD can easily extract patterns, but the user needs to know in advance what he is looking for. MOD systems are not good for data exploration. As a future work, we plan to break the direct manipulation data set limitation with new interaction paradigms (more complex Boolean operations). We also plan to combine MOD, Visual Analytics, and direct manipulation to explore large data sets. Visualize a small sample, roughly figure out your query parameters, issue the query in MOD, validate the results by visual analytics, refine the MOD query, and so on and so forth.

12.8 Bibliographic Notes

For further reading, we recommend the book by Card et al. (1999) which details the information visualization research area. We also recommend the book by Tufte (1990) which contains many remarkable visualization instances. Two conference proceedings contain many examples of visualizations and interaction techniques. InfoVis: The IEEE Information Visualization Conference (IEEE Transactions on Visualization and Computer Graphics) contains novel research ideas and innovative applications in all areas of information visualization. Also, VAST, the IEEE Conference on Visual Analytics Science and Technology is the first international conference dedicated to advances in Visual Analytics Science and Technology. The scope of the conference includes both fundamental research contributions within visual analytics as well as applications of visual analytics, including applications in science, engineering, medicine, health, media, business, social interaction, and security and investigative analysis.

The spatio-temporal pattern predicate was first proposed in Sakr and Güting

(2011). It is demonstrated in Sakr et al. (2011). We used this demonstration as the basis of Section 12.6.

Bibliography

Abul, O., Bonchi, F., and Giannotti, F. 2010. Hiding sequential and spatiotemporal patterns. *IEEE Transactions on Knowledge and Data Engineering*, **22**(12), 1709–1723.

Agrawal, R., and Srikant, R. 2002. Mining sequential patterns. Pages 3–14 of: *Proceedings of the 11th International Conference on Data Engineering*. IEEE.

Ahas, R., Aasa, A., Roose, A., Mark, Ü., and Silm, S. 2008. Evaluating passive mobile positioning data for tourism surveys: An Estonian case study. *Tourism Management*, **29**(3), 469–486.

Alt, H., and Godau, M. 1995. Computing the Fréchet distance between two polygonal curves. *International Journal of Computational Geometry and Applications*, **5**(1), 75–91.

Alt, W., and G., Hoffman. 1990. *Biological Motion*. Springer-Verlag.

Alvares, L.O., Bogorny, V., Kuijpers, B., Macedo, J.A., Moelans, B., and Vaisman, A. 2007. A model for enriching trajectories with semantic geographical information. Pages 1–8 of: *Proceedings of the 15th Annual ACM International Symposium on Advances in Geographic Information Systems*. ACM Press.

Andrienko, G., and Andrienko, N. 2010. A general framework for using aggregation in visual exploration of movement data. *The Cartographic Journal*, **47**(1), 22–40.

Andrienko, G., Andrienko, N., Bak, P., Keim, D., Kisilevich, S., and Wrobel, S. 2011a. A conceptual framework and taxonomy of techniques for analyzing movement. *Journal of Visual Languages & Computing*, **22**(3), 213–232.

Andrienko, G., Andrienko, N., and Heurich, M. 2011b. An event-based conceptual model for context-aware movement analysis. *International Journal of Geographical Information Science*, **25**(9), 1347–1370.

Andrienko, G., Andrienko, N., Hurter, C., Rinzivillo, S., and Wrobel, S. 2011c. From movement tracks through events to places: Extracting and characterizing significant places from mobility data. Pages 161–170 of: *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*. IEEE Computer Society Press.

Andrienko, N., and Andrienko, G. 2011. Spatial generalization and aggregation of massive movement data. *IEEE Transactions on Visualization and Computer Graphics*, **17**(2), 205–219.

Andrienko, N., and Andrienko, G. 2012. Visual analytics of movement: An overview of methods, tools and procedures. *Information Visualization*. To appear.

Ankerst, M., Breunig, M.M., Kriegel, H.-P., and Sander, J. 1999. OPTICS: Ordering points to identify the clustering structure. Pages 49–60 of: *Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data*. ACM Press.

Baglioni, M., Macedo, J.A., Renso, C., Trasarti, R., and Wachowicz, M. 2012. How you move reveals who you are: Understanding human behavior by analyzing trajectory data. *Knowledge and Information System Journal*. Online first.

Barabási, A. L., and Albert, R. 1999. Emergence of Scaling in Random Networks. *Science*, **286**, 509.

Barrett, G. 1996. The transport dimension. In: *The Compact City: A Sustainable Urban Form?* E & FN Spon.

Bertrand, F., Bouju, A., Claramunt, C., T., Devogele, and Ray, C. 2007. Web architecture for monitoring and visualizing mobile objects in maritime contexts. Pages 94–105 of: *Proceedings of the 7th International Symposium on Web and Wireless Geographical Information Systems*. LNCS 4857. Springer-Verlag.

Bole, A., Dineley, W., and Wall, A. 2012. *Radar and ARPA Manual: Radar and Target Tracking for Professional Mariners, Yachtsmen and Users of Marine Radar*. Third edn. Butterworth-Heinemann.

Bonchi, F., Lakshmanan, L.V.S., and Wang, W.H. 2011. Trajectory anonymity in publishing personal mobility data. *SIGKDD Explorations*, **13**(1), 30–42.

Bouvier, D.J., and Oates, B. 2008. Evacuation traces mini challenge award: Innovative trace visualization. Staining for information discovery. Pages 219–220 of: *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology*. IEEE Computer Society Press.

Brakatsoulas, S., Pfoser, D., Salas, R., and Wenk, C. 2005. On map-matching vehicle tracking data. Pages 853–864 of: *Proceedings of the 31st International Conference on Very Large Data Bases*. VLDB Endowment.

Bremm, S., Andrienko, G., Andrienko, N., Schreck, T., and Landesberger, T. 2011. Interactive analysis of object group changes over time. Pages 41–44 of: *Proceedings of the International Workshop on Visual Analytics*. Eurographics.

Brockmann, D., Hufnagel, L., and Geisel, T. 2006. The scaling laws of human travel. *Nature*, **439**, 462–465.

Bruzzese, D., and Davino, C. 2003. Visual post-analysis of association rules. *Journal of Visual Languages & Computing*, **14**(6), 621–635.

Burke, R.R. 2005. The third wave of marketing intelligence. Pages 103–115 of: Krafft, M., and Mantrala, M.K (eds), *Retailing in the 21st Century*. Springer-Verlag.

Cabibbo, L., and Torlone, R. 1997. Querying multidimensional databases. Pages 319–335 of: *Proceedings of the 6th International Workshop on Database Programming Languages*. LNCS 1396. Springer-Verlag.

Cagnacci, F., Boitani, L., Powell, R.A., and Boyce, M.S. (eds). 2010. *Philosophical Transactions of the Royal Society B, volume 365*. Special issue on Challenges and opportunities of using GPS location data in animal ecology.

Cao, H., Mamoulis, N., and Cheung, D.W. 2005. Mining frequent spatio-temporal sequential patterns. Pages 82–89 of: *Proceedings of the 5th International Conference on Data Mining*. IEEE Computer Society Press.

Card, S.K., Mackinlay, J.D., and Shneiderman, B. (eds). 1999. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann.

Cho, E., Myers, S. A., and Leskovec, J. 2011a. Friendship and mobility: User movement in location-based social networks. Pages 1082–1090 of: *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM Press.

Cho, E., Myers, S.A., and Leskovec, J. 2011b. Friendship and mobility: user movement in location-based social networks. Pages 1082–1090 of: *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM Press.

Chorley, M.J., Colombo, G.B., Williams, M.J., Allen, S.M., and Whitaker, R.M. 2011. Checking out checking in: observation on Foursquare usage patterns. Pages 28–39 of: *Proceedings of the International Workshop on Finding Patterns of Human Behaviors in Networks and Mobility Data*.

Chow, C., Mokbel, M.F., and Aref, W.G. 2009. Casper*: Query processing for location services without compromising privacy. *ACM Transactions on Database Systems*, **34**(4), 24.

Coscia, M., Giannotti, F., and Pedreschi, D. 2011. A classification for community discovery methods in complex networks. *Statistical Analysis and Data Mining*, **4**(5), 512–546.

Craglia, M., Ostermann, F., and Spinsanti, L. 2012. Digital Earth from vision to practice: making sense of citizen-generated content. *International Journal of Digital Earth*, **5**(5), 398–416.

Crandall, D. J., Backstrom, L., Cosley, D., Suri, S., Huttenlocher, D., and Kleinberg, J. 2010. Inferring social ties from geographic coincidences. *Proceedings of the National Academy of Sciences*.

Cranshaw, J., Toch, E., Hong, J., Kittur, A., and Sadeh, N. 2010. Bridging the gap between physical location and online social networks. Pages 119–128 of: *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*. ACM Press.

Damiani, M.L., Bertino, E., and Silvestri, C. 2010. The PROBE framework for the personalized cloaking of private locations. *Transactions on Data Privacy*, **(3)2**, 123–148.

Damiani, M.L., Silvestri, C., and Bertino, E. 2011. Fine-grained cloaking of sensitive positions in location-sharing applications. *IEEE Pervasive Computing*, **10**(4), 64–72.

Dee, H.M., and Velastin, S.A. 2007. How close are we to solving the problem of automated visual surveillance? *Machine Vision and Applications*, **19**(5-6), 329–343.

Devogele, T. 2002. A new merging process for data integration based on the discrete Fréchet distance. Pages 167–181 of: *Proceedings of the 10th International Symposium on Spatial Data Handling*. Springer-Verlag.

Dodge, S., Weibel, R., and Lautenschütz, A.-K. 2008. Taking a systematic look at movement: Developing a taxonomy of movement patterns. In: *Proceedings of the AGILE Workshop on GeoVisualization of Dynamics, Movement and Change*.

Domingo-Ferrer, J., and Trujillo-Rasua, R. 2012. Microaggregation- and permutation-based anonymization of movement data. *Information Sciences*, **208**, 55–80.

Dricot, J.-M., Bontempi, G., and Doncker, P. 2009. Static and dynamic localization techniques for wireless sensor networks. Pages 249–281 of: Ferrari, G. (ed), *Sensor Networks: Where Theory Meets Practice*. Springer-Verlag.

Düntgen, C., Behr, T., and Güting, R.H. 2009. BerlinMOD: A benchmark for moving object databases. *VLDB Journal*, **18**(6), 1335–1368.

Eagle, N., and Pentland, A. 2005. Reality mining: Sensing complex social systems. *Personal and Ubiquitous Computing*, **10**(4), 255–268.

Egenhofer, M. 2003. Approximation of geospatial lifelines. In: *Proceedings of the Workshop on Spatial Data and Geographic Information Systems*.

Erdös, P., and Rényi, A. 1959. On random graphs. *Publicationes Mathematicae (Debrecen)*, **6**, 290–297.

Etienne, L., Devogele, T., and Bouju, A. 2012. Spatio-temporal trajectory analysis of mobile objects following the same itinerary. Pages 47–58 of: Shi, W., Goodchild, M., Lees, B., and Leung, Y. (eds), *Advances in Geo-Spatial Information Science*. CRC Press.

Fayyad, U.M., Piatetsky-Shapiro, G., and Smyth, P. 1996. From data mining to knowledge discovery: An overview. Pages 1–34 of: *Advances in Knowledge Discovery and Data Mining*. American Association for Artificial Intelligence.

Focardi, S., Montanaro, P., and Pecchioli, E. 2009. Adaptive Lévy walks in foraging fallow deer. *PloS ONE*. doi 10.1371/journal.pone.0006587.

Fraenkel, G.S., and Gunn, D.L. 1961. *The Orientation of Animals: Kineses, Taxes, and Compass Reactions*. Dover Publications.

Fryxell, J.M., Hazell, M., Börger, L., Dalziel, B.D., Haydon, D.T., Morales, J.M., McIntosh, T., and Rosatte, R.C. 2008. Multiple movement modes by large herbivores at multiple spatio-temporal scales. *Proceedings of the National Academy of Sciences*, **105**, 19114–19119.

Gaffney, S., and Smyth, P. 1999. Trajectory clustering with mixture of regression models. Pages 63–72 of: *Proceedings of the 5th International Conference on Knowledge Discovery and Data Mining*. ACM Press.

Gaito, S., Rossi, G.P., and Zignani, M. 2011. From mobility data to social attitudes: a complex network approach. Pages 52–65 of: *Proceedings of the International Workshop on Finding Patterns of Human Behaviors in Networks and Mobility Data*.

Ghinita, G., Damiani, M.L., Silvestri, C., and Bertino, E. 2009. Preventing velocity-based linkage attacks in location-aware applications. Pages 246–255 of: *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM Press.

Giannotti, F., and Pedreschi, D. (eds). 2008. *Mobility, Data Mining and Privacy: Geographic Knowledge Discovery*. Springer-Verlag.

Giannotti, F., Nanni, M., Pinelli, F., and Pedreschi, D. 2007. Trajectory pattern mining. Pages 330–339 of: *Proceedings of the 13th International Conference on Knowledge Discovery and Data Mining*. ACM Press.

Giannotti, F., Nanni, M., Pedreschi, D., Pinelli, F., Renso, C., Rinzivillo, S., and Trasarti, R. 2011. Unveiling the complexity of human mobility by querying and mining massive trajectory data. *VLDB Journal*, **20**(5), 695–719.

Gómez, L.I., Gómez, S., and Vaisman, A. 2012. A Generic Data Model and Query Language for Spatiotemporal OLAP Cube Analysis. Pages 300–311 of: *Proceedings of the 15th International Conference on Extending Database Technology*. ACM Press.

González, M. C., Hidalgo, C. A., and Barabási, A. L. 2008. Understanding human mobility patterns. *Nature*, **454**, 779–782.

Gould, J.L., and Gould, C.G. 2012. *Nature's Compass: The Mystery of Animal Navigation*. Princeton University Press.

Granovetter, M. S. 1973. The strength of weak ties. *America Journal of Sociology*, **78**, 1360–1380.

Greenfeld, Joshua S. 2002. Matching GPS observations to locations on a digital map. *Proceedings of the 81st Annual Meeting of the Transportation Research Board*, **1**(3), 164–173.

Gruteser, M., and Grunwald, D. 2003. Anonymous usage of location-based services through spatial and temporal cloaking. Pages 31–42 of: *Proceedings of the 1st International Conference on Mobile Systems, Applications and Services*. ACM Press.

Gudmundsson, J., van Kreveld, M.J., and Speckmann, B. 2004. Efficient detection of motion patterns in spatio-temporal data sets. Pages 250–257 of: *Proceedings of the 12th International Workshop on Geographic Information Systems*. ACM Press.

Guo, D. 2007. Visual analytics of spatial interaction patterns for pandemic decision support. *International Journal of Geographical Information Science*, **21**(8), 859–877.

Guo, H., Wang, Z., Yu, B., Zhao, H., and Yuan, X. 2011. TripVista: Triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection. Pages 163–170 of: *Proceedings of the IEEE Pacific Visualization Symposium*. IEEE Computer Society Press.

Güting, R.H., and Schneider, M. 2005. *Moving Objects Databases*. Morgan Kaufmann.

Güting, R.H., Böhlen, M.H., Erwig, M., Jensen, C.S., Lorentzos, N.A., Schneider, M., and Vazirgiannis, M. 2000. A foundation for representing and querying moving objects. *ACM Transactions on Database Systems*, **25**(1), 1–42.

Güting, R.H., Teixeira de Almeida, V., and Ding, Z. 2006. Modeling and querying moving objects in networks. *VLDB Journal*, **15**(2), 165–190.

Güting, R.H., Behr, T., and Xu, J. 2010. Efficient *k*-nearest neighbor search on moving object trajectories. *VLDB Journal*, **19**(5), 687–714.

Hägerstrand, T.H. 1970. What about people in regional science? *Papers of the Regional Science Association*, **24**, 7–21.

Haghani, A., Hamedi, M., Sadabadi, K.F., Young, S., and Tarnoff, P. 2010. Data collection of freeway travel time ground truth with Bluetooth sensors. *Transportation Research Record: Journal of the Transportation Research Board*, **2160**, 60–68.

Hurter, C., Tissoires, B., and Conversy, S. 2009. FromDaDy: Spreading aircraft trajectories across views to support iterative queries. *IEEE Transactions on Visualization and Computer Graphics*, **15**(6), 1017–1024.

IALA, International Association of Marine Aids to Navigation & Lighthouse Authorities. 2004. *The Automatic Identification System (AIS), Volume 1, Part I, Operational Issues, Edition 1.3*. IALA Guideline No. 1028.

IHO, International Hydrographic Organization. 2000. *Transfer Standard for Digital Hydrographic Data, Edition 3.1*. Special Publication No. 57.

IMO, International Maritime Organization. 2008. *Development of an E-Navigation Strategy*. Reports of Sub-Committee on Safety of Navigation, 54th session.

Inan, A., and Saygin, Y. 2006. Privacy preserving spatio-temporal clustering on horizontally partitioned data. Pages 459–468 of: *Proceedings of the 8th International Conference on Data Warehousing and Knowledge Discovery*. LNCS 4081. Springer-Verlag.

Jankowski, P., Andrienko, N., Andrienko, G., and Kisilevich, S. 2011. Discovering landmark preferences and movement patterns from photo postings. *Transactions in GIS*, **14**(6), 833–852.

Jensen, C.S., Lu, H., and Yiu, M.L. 2009. Location privacy techniques in client-server architectures. Pages 31–58 of: Bettini, C., Jajodia, S., Samarati, P., and Wang, S.X. (eds), *Privacy in Location-Based Applications: Research Issues and Emerging Trends*. Springer-Verlag.

Jeung, H., Yiu, M.L., Zhou, X., Jensen, C.S., and Shen, H.T. 2008. Discovery of convoys in trajectory databases. *Proceedings of the VLDB*, **1**(1), 1068–1080.

Johnson, D. 1980. The comparison of usage and availability measurements for evaluating resource preference. *Ecology*, **61**, 65–71.

Kalnis, P., Mamoulis, N., and Bakiras, S. 2005. On discovering moving clusters in spatio-temporal data. Pages 364–381 of: *Proceedings of the 9th International Symposium on Spatial and Temporal Databases*. LNCS 3633. Springer-Verlag.

Kapler, T., and Wright, W. 2005. GeoTime information visualization. *Information Visualization*, **4**(2), 136–146.

Karamshuk, D., Boldrini, C., Conti, M., and Passarella, A. 2011. Human mobility models for opportunistic networks. *IEEE Communication Magazine*, **49**(12), 157–165.

Keim, D., Andrienko, G., Fekete, J.-D., Carsten, G., Kohlhammer, J., and Melançon, G. 2008. Visual Analytics: Definition, Process, and Challenges. Chap. 7, pages 154–175 of: Kerren, A., Stasko, J., Fekete, J.-D., and North, C. (eds), *Information Visualization*. LNCS 4950. Springer-Verlag.

Kellaris, G., Pelekis, N., and Theodoridis, Y. 2009. Trajectory compression under network constraints. Pages 392–398 of: *Proceedings of the 11th International Symposium on Advances in Spatial and Temporal Databases*. LNCS 5644. Springer-Verlag.

Kimball, R. 1996. *The Data Warehouse Toolkit*. J. Wiley and Sons.

Kisilevich, S., Keim, D.A., and Rokach, L. 2010a. A novel approach to mining travel sequences using collections of geotagged photos. Pages 163–182 of: *Proceedings of the 13th AGILE International Conference on Geographic Information Science*. Lecture Notes in Geoinformation and Cartography. Springer-Verlag.

Kisilevich, S., Mansmann, F., Nanni, M., and Rinzivillo, S. 2010b. Spatio-temporal clustering. Pages 855–874 of: Maimon, O., and Rokach, L. (eds), *Data Mining and Knowledge Discovery Handbook*, second edn. Springer-Verlag.

Klug, A. 1982. Equivalence of relational algebra and relational calculus query languages having aggregate functions. *Journal of the ACM*, **29**(3), 699–717.

Koubarakis, M., Sellis, T., Frank, A., Guting, R., Jensen, C.S., Lorentzos, A., Manolopoulos, Y., Nardelli, E., Pernici, B., Schek, H.-J., Scholl, M., Theodoulidis, B., and Tryfona, N. 2003. *Spatio-Temporal Databases: The Chorochronos Approach*. Springer-Verlag.

Kuijpers, B., and Othman, W. 2009. Modeling uncertainty of moving objects on road networks via space-time prisms. *International Journal of Geographical Information Science*, **23**(9), 1095–1117.

Kuijpers, B., Moelans, B., Othman, W., and Vaisman, A.A. 2009. Analyzing trajectories using uncertainty and background information. Pages 135–152 of: *Proceedings of the 11th International Symposium on Spatial and Temporal Databases*. LNCS 5644. Springer-Verlag.

Laube, P., van Kreveld, M., and Imfeld, S. 2005. Finding REMO - Detecting Relative Motion Patterns in Geospatial Lifelines. Pages 201–214 of: *Proceedings of the 11th International Symposium on Spatial Data Handling*. Springer-Verlag.

Lee, J.-G., Han, J., Li, X., and Gonzalez, H. 2008a. TraClass: Trajectory classification using hierarchical region-based and trajectory-based clustering. *Proceedings of the VLDB Endowment*, **1**(1), 1081–1094.

Lee, J.-G., Han, J., and Li, X. 2008b. Trajectory outlier detection: A partition-and-detect framework. Pages 140–149 of: *Proceedings of the 24th International Conference on Data Engineering*. IEEE Computer Society Press.

Leitinger, S., Gröchenig, S., Pavelka, S., and Wimmer, M. 2010. Erfassung von personenströmen mit der Bluetooth-tracking technologie. Pages 220–225 of: *Angewandte Geoinformatik 2010*.

Li, Ninghui, Li, Tiancheng, and Venkatasubramanian, Suresh. 2007. t-Closeness: Privacy beyond k-anonymity and l-diversity. Pages 106–115 of: *Proceedings of the IEEE 23rd International Conference on Data Engineering*. IEEE Computer Society Press.

Lucchese, C., Perego, R., Silvestri, F., Vahabi, H., and Venturini, R. 2012. How random walks can help tourism. Pages 195–206 of: *Proceedings of the 34th European conference on Advances in Information Retrieval (ECIR'12)*. Lecture Notes in Computer Science. Springer-Verlag.

Machanavajjhala, A., Kifer, D., Gehrke, J., and Venkitasubramaniam, M. 2007. *L*-diversity: Privacy beyond *k*-anonymity. *ACM Transactions on Knowledge Discovery from Data*, **1**(1), 3.

Malinowski, E., and Zimányi, E. 2008. *Advanced Data Warehouse Design: From Conventional to Spatial and Temporal Applications*. Springer-Verlag.

Mannila, H. 1997. Inductive Databases and Condensed Representations for Data Mining. Pages 21–30 of: *Proceedings of the 1997 International Symposium on Logic Programming*. MIT Press.

Marketos, G., Frentzos, E., Ntoutsi, I., Pelekis, N., Raffaetà, A., and Theodoridis, Y. 2008. Building real-world trajectory warehouses. Pages 8–15 of: *Proceedings of the 7th ACM International Workshop on Data Engineering for Wireless and Mobile Access*. ACM Press.

MarNIS, Maritime Navigation Information Services. 2009. *Final Report*. MarNIS/D-MT-15/Final Report/DVS/05062009/version 2.0.

Meratnia, N., and de By, R.A. 2004. Spatio-temporal compression techniques for moving point objects. Pages 765–782 of: *Proceedings of the 9th International Conference on Extending Database Technology*. LNCS 2992. Springer-Verlag.

Milgram, S. 1967. The small world problem. *Psychology Today*, **2**, 60–67.

Miller, H.J. 1991. Modeling accessibility using space-time prism concepts within geographical information systems. *International Journal of Geographical Information Systems*, **5**(3), 287–301.

Miller, H.J. 2010. The data avalanche is here. Shouldn't we be digging? *Journal of Regional Science*, **50**(1), 181–201.

Miller, N.P. 2003. Transportation noise and recreational lands. *Noise/News International*, **11**(2), 9–20.

Mlich, J., and Chmelar, P. 2008. Trajectory classification based on hidden Markov models. Pages 101–105 of: *Proceedings of the 18th International Conference on Computer Graphics and Vision*.

Mokbel, M.F., Ghanem, T.M., and Aref, W.G. 2003. Spatio-temporal access methods. *IEEE Data Engineering Bulletin*, **26**(2), 40–49.

Monreale, A. 2011. *Privacy by Design in Data Mining*. Ph.D. thesis, Department of Computer Science, University of Pisa, Italy.

Monreale, A., Pinelli, F., Trasarti, R., and Giannotti, F. 2009. WhereNext: A location predictor on trajectory pattern mining. Pages 637–646 of: *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM Press.

Monreale, A., Pedreschi, D., and Pensa, R.G. 2010. Anonymity technologies for privacy-preserving data publishing and mining. Pages 3–33 of: Bonchi, F., and Ferrari, E. (eds), *Privacy-Aware Knowledge Discovery: Novel Applications and New Techniques*. Chapman & Hall/CRC Press.

Moore, B.E., Ali, S., Mehran, R., and Shah, M. 2011. Visual crowd surveillance through a hydrodynamics lens. *Communications of the ACM*, **54**(12), 64–73.

Nanni, M., and Pedreschi, D. 2006. Time-focused clustering of trajectories of moving objects. *Journal of Intelligent Information Systems*, **27**(3), 267–289.

Nathan, R., Getz, W.M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., and Smouse, P.E. 2008. A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences*, **105**, 19052–19059.

Newson, P., and Krumm, J. 2009. Hidden Markov map matching through noise and sparseness. Pages 336–343 of: *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM Press.

Nguyen-Dinh, L.-V., Aref, W.G., and Mokbel, M.F. 2010. Spatio-temporal access methods: Part 2 (2003 - 2010). *IEEE Data Engineering Bulletin*, **33**(2), 46–55.

Okubo, A. 1980. *Diffusion and ecological problems: mathematical models*. Springer-Verlag.

Orlando, S., Orsini, R., Raffaetà, A., Roncato, A., and Silvestri, C. 2007. Spatio-temporal aggregations in trajectory data warehouses. *Journal of Computing Science and Engineering*, **1**(2), 211–232.

Ortúzar, J., and Willumsen, L.G. 2002. *Modelling Transport*. John Wiley and Sons.

Othman, W. 2009. *Uncertainty Management in Trajectory Databases*. Ph.D. thesis, Transnationale Universiteit Limburg.

Parent, C., Spaccapietra, S., and Zimányi, E. 2006. *Conceptual Modeling for Traditional and Spatio-Temporal Applications: The MADS Approach*. Springer-Verlag.

Pauly, A., and Schneider, M. 2010. VASA: An algebra for vague spatial data in databases. *Information Systems*, **35**(1), 111–138.

Pelekis, N., and Theodoridis, Y. 2005. *An Oracle Data Cartridge for Moving Objects*. Tech. rept. University of Peraeus.

Pelekis, N., Kopanakis, I., Marketos, G., Ntoutsi, I., Andrienko, G., and Theodoridis, Y. 2007. Similarity search in trajectory databases. Pages 129–140 of: *Proceedings of the 14th International Symposium on Temporal Representation and Reasoning*. IEEE Computer Society.

Pelekis, N., Frentzos, E., Giatrakos, N., and Theodoridis, Y. 2008a. HERMES: Aggregative LBS via a trajectory DB engine. Pages 1255–1258 of: *Proceedings of the ACM SIGMOD Conference on Management of Data*. ACM Press.

Pelekis, N., Raffaetà, A., Damiani, M. L., Vangenot, C., Marketos, G., Frentzos, E., Ntoutsi, I., and Theodoridis, Y. 2008b. Towards trajectory data warehouses. In: Giannotti, F., and Pedreschi, D. (eds), *Mobility, Data Mining and Privacy*. Springer-Verlag.

Pelekis, N., Kopanakis, I., Kotsifakos, E.E., Frentzos, E., and Theodoridis, Y. 2011. Clustering uncertain trajectories. *Knowledge and Information Systems*, **28**(1), 117–147.

Peterson, B.S., Baldwin, R.O., and Kharoufeh, J.P. 2006. Bluetooth inquiry time characterization and selection. *IEEE Transactions on Mobile Computing*, **5**(9), 1173–1187.

Pfoser, D., and Jensen, C.S. 1999. Capturing the uncertainty of moving-object representations. Pages 111–132 of: *Proceedings of the 6th International Symposium on Advances in Spatial Databases*. LNCS 1651. Springer-Verlag.

Pfoser, D., Jensen, C.S., and Theodoridis, Y. 2000. Novel approaches in query processing for moving object trajectories. Pages 395–406 of: *Proceedings of the 26th International Conference on Very Large Data Bases*. Morgan Kaufmann.

Potamias, M., Patroumpas, K., and Sellis, T. 2006. Sampling trajectory streams with spatiotemporal criteria. Pages 275–284 of: *Proceedings of the 18th International Conference on Scientific and Statistical Database Management*. IEEE Computer Society.

Purves, R.S. 2011. Answering geographic questions with user generated content: experiences from the coal face. Pages 297–299 of: *Proceedings of the 27th Annual ACM Symposium on Computational Geometry*. ACM Press.

Quddus, M.A., Ochieng, W.Y., and Noland, R.B. 2007. Current map-matching algorithms for transport applications: State-of-the art and future research directions. *Transportation Research Part C: Emerging Technologies*, **15**(5), 312–328.

Raffaetà, A., Leonardi, L., Marketos, G., Andrienko, G., Andrienko, N., Frentzos, E., Giatrakos, N., Orlando, S., Pelekis, N., Roncato, A., and Silvestri, C. 2011. Visual mobility analysis using T-Warehouse. *International Journal of Data Warehousing and Mining*, **7**(1), 1–23.

Ratti, C., Sobolevsky, S., Calabrese, F., Andris, C., Reades, J., Martino, M., Claxton, R., and Strogatz, S.H. 2010. Redrawing the map of Great Britain from a network of human interactions. *PLoS One*, **5**(12), e14248+.

Rietveld, P. 1994. Spatial economic impacts of transport infrastructure supply. *Transportation Research-A*, **28A**, 329–341.

Rinzivillo, S., Pedreschi, D., Nanni, M., Giannotti, F., Andrienko, N., and Andrienko, G. 2008. Visually driven analysis of movement data by progressive clustering. *Information Visualization*, **7**(3–4), 225–239.

Rinzivillo, S., Mainardi, S., Pezzoni, F., Coscia, M., Pedreschi, D., and Giannotti, F. 2012. Discovering the geographical borders of human mobility. *Künstliche Intelligenz*, **26**(3), 253–260.

Ruiter, E.R., and Ben-Akiva, M.E. 1978. Disaggregate travel demand models for the San Francisco Bay Area. *Transportation Research Record*, **673**, 121–128.

Sakoe, H., and Chiba, S. 1978. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **26**(1), 43–49.

Sakr, M., Güting, R.H., Behr, T., Adrienko, G., Andrienko, N., and Hurter, C. 2011. Exploring spatiotemporal patterns by integrating visual analytics with a moving objects database system. Pages 505–508 of: *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM Press.

Sakr, M.A., and Güting, R.H. 2011. Spatiotemporal pattern queries. *GeoInformatica*, **15**(3), 497–540.

Salomon, I., Bovy, P., and J.-P., Orfeuil. 1993. *A Billion Trips a Day*. Kluwer.

Samarati, P., and Sweeney, L. 1998. Generalizing data to provide anonymity when disclosing information (abstract). Page 188 of: *Proceedings of the 17th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*. ACM Press.

Scheaffer, R.L., W. Mendenhall, III, Ott, R.L., and Gerow, K. 2005. *Elementary Survey Sampling*. Richard Stratton.

Shneiderman, B. 1983. Direct Manipulation: A Step Beyond Programming Languages. *Computer*, **16**(8), 57–69.

Shu, H., Spaccapietra, S., and Quesada Sedas, D. 2003. Uncertainty of geographic information and its support in MADS. In: *Proceedings of the 2nd International Symposium on Spatial Data Quality*.

Smouse, P.E., Focardi, S., Moorcroft, P.R., Kie, J.G., Forester, J.D., and Morales, J.M. 2010. Stochastic modelling of animal movement. *Philosophical Transactions of the Royal Society B*, **365**, 2201–2211.

Song, C., Qu, Z., Blumm, N., and Barabási, A. L. 2009. Limits of Predictability in Human Mobility. *Science*, **327**, 1018–1021.

Song, C., Koren, T., Wang, P., and Barabási, A. L. 2010. Modelling the scaling properties of human mobility. *Nature Physics*, **6**, 818–823.

Spaccapietra, S., Parent, C., Damiani, M.L., Macedo, J.A., Porto, F., and Vangenot, C. 2008. A conceptual view on trajectories. *Data & Knowledge Engineering*, **65**(1), 126–146.

Stange, H., Liebig, T., Hecker, D., Andrienko, G., and Andrienko, N. 2011. Analytical workflow of monitoring human mobility in big event settings using Bluetooth. Pages 51–58 of: *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness*. ACM Press.

Tan, P.-N., Steinbach, M., and Kumar, V. 2005. *Introduction to Data Mining*. Addison-Wesley.

Teixeira de Almeida, V., and Güting, R.H. 2005. Indexing the trajectories of moving objects in networks. *GeoInformatica*, **9**(1), 33–60.

Tobler, W.R. 1987. Experiments in migration mapping by computer. *The American Cartographer*, **14**(2), 155–163.

Tomkiewicz, S.M., Fuller, M.R., Kie, J.G., and Bates, K.K. 2010. Global positioning system and associated technologies in animal behaviour and ecological research. *Philosophical Transaction Royal Society B*, **365**, 2163–2176.

Trajcevski, G. 2011. Uncertainty in spatial trajectories. Pages 63–107 of: Zheng, Y., and Zhou, X. (eds), *Computing with Spatial Trajectories*. Springer-Verlag.

Trajcevski, G., Wolfson, O., Hinrichs, K., and Chamberlain, S. 2004. Managing uncertainty in moving objects databases. *ACM Transactions on Database Systems*, **29**(3), 463–507.

Trasarti, R., Giannotti, F., Nanni, M., Pedreschi, D., and Renso, C. 2011. A query language for mobility data mining. *International Journal of Data Warehouse and Mining*, **7**(1), 24–45.

Tufte, E. 1990. *Envisioning Information*. Graphics Press.

Tukey, J.W. 1977. *Exploratory Data Analysis*. Addison-Wesley.

Turchin, P. 1998. *Quantitative analysis of movement*. Sinauer Associates, Inc.

Urbano, F., Cagnacci, F., Calenge, C., Cameron, A., and Neteler, M. 2010. Wildlife tracking data management: a new vision. *Philosophical Transactions of the Royal Society*, **365**, 2177–2186.

Vaisman, A., and Zimányi, E. 2009a. A multidimensional model representing continuous fields in spatial data warehouses. Pages 168–177 of: *Proceedings of the 17th ACM SIGSPATIAL International Symposium on Advances in Geographic Information Systems*. ACM Press.

Vaisman, A., and Zimányi, E. 2009b. What is spatio-temporal data warehousing? Pages 9–23 of: *Proceedings of the 11th International Conference on Data Warehousing and Knowledge Discovery*. LNCS 5691. Springer-Verlag.

Van der Spek, S., Van Schaick, J., De Bois, P., and De Haan, R. 2009. Sensing human activity: GPS tracking. *Sensors*, **9**(4), 3033–3055.

Versichele, M., Neutens, T., Delafontaine, M., and Van de Weghe, N. 2012. The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the Ghent festivities. *Applied Geography*, **32**(2), 208–220.

Viswanathan, G.M., da Luz, M.G.E., Raposo, E.P., and Stanley, H.E. 2011. *The Physics of Foraging*. Cambridge University Press.

Wang, D., Pedreschi, D., Song, C., Giannotti, F., and Barabási, A. L. 2011. Human mobility, social ties, and link prediction. Pages 1100–1108 of: *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM Press.

Warf, B., and Sui, D. 2010. From GIS to neogeography: ontological implications and theories of truth. *Annals of GIS*, **16**(4), 197–209.

Watts, D. J., and Strogatz, S. H. 1998. Collective dynamics of 'small-world' networks. *Nature*, **393**, 440.

Willems, N., Van De Wetering, H., and Van Wijk, J.J. 2009. Visualization of vessel movements. *Computer Graphics Forum*, **28**(3), 959–966.

Wood, Z., and Galton, A. 2009. Classifying Collective Motion. Pages 129–155 of: Gottfried, B., and Aghajan, H. (eds), *Behaviour Monitoring and Interpretation - BMI: Smart Environments*. IOS Press.

Wood, Z., and Galton, A. 2010. Zooming in on collective motion. Pages 25–30 of: *Proceedings of the ECAI 2010 Workshop on Spatio-Temporal Dynamics*.

Xu, J., and Güting, R.H. 2012. A generic data model for moving objects. *GeoInformatica*. Online first.

Yan, Z., Parent, C., S., Stefano, and Chakraborty, D. 2010. A hybrid model and computing platform for spatio-semantic trajectories. Pages 60–75 of: *Proceedings of the 7th International Conference on The Semantic Web: Research and Applications - Volume Part I*. LNCS 6088. Springer-Verlag.

Yan, Z., Chakraborty, D., Parent, C., Spaccapietra, S., and Aberer, K. 2011. SeMiTri: A framework for semantic annotation of heterogeneous trajectories. Pages 259–270 of: *Proceedings of the 14th International Conference on Extending Database Technology*. ACM Press.