# Visual Analytics of Mobility and Transportation: State of the Art and Further Research Directions

Gennady Andrienko, Natalia Andrienko, Wei Chen, Ross Maciejewski, and Ye Zhao

***Abstract*—Many cities and countries are now striving to create intelligent transportation systems that utilize the current abundance of multisource and multiform data related to the functionality and use of transportation infrastructure to better support human mobility, interests, and lifestyles. Such intelligent transportation systems aim to provide novel services that can enable transportation consumers and managers to be better informed and make safer and more efficient use of the infrastructure. However, the transportation domain is characterized by both complex data and complex problems, which calls for visual analytics approaches. The science of visual analytics is continuing to develop principles, methods, and tools to enable synergistic work between humans and computers through interactive visual interfaces. Such interfaces support the unique capabilities of humans (such as the flexible application of prior knowledge and experiences, creative thinking, and insight) and couple these abilities with machines' computational strengths, enabling the generation of new knowledge from large and complex data. In this paper, we describe recent developments in visual analytics that are related to the study of movement and transportation systems and discuss how visual analytics can enable and improve the intelligent transportation systems of the future. We provide a survey of literature from the visual analytics domain and organize the survey with respect to different types of transportation data, movement and its relationship to infrastructure and behavior, and modeling and planning. We conclude with lessons learned and future directions including social transportation, recommender systems, and policy implications.**

***Index Terms*—Data visualization, graphical user interfaces, interactive systems**

This work was supported in part by EU in projects datAcron (grant agreement 687591), VaVeL (grant agreement 688380), DART (grant agreement 699299), INTUIT (grant agreement 699303), and BigData4ATM (grant agreement 699260). Chen's work was partly support by National 973 Program of China (2015CB352503) and NSFC (No. U1609217). Maciejewski's work was partially supported by the National Science Foundation under grant number 1350573. Zhao's work is partly supported by National Science Foundation grants 1535031 and 1637242.

G.Andrienko and N.Andrienko are with Fraunhofer Institute IAIS (Sankt Augustin, Germany) and City University London, UK (e-mail: {Gennady|Natalia}.Andrienko@iais.fraunhofer.de).

W.Chen is with Zhejiang University, China (e-mail: chenwei@cad.zju.edu.cn).

R.Maciejewski is with Arizona State University, USA (e-mail: rmacieje@asu.edu).

Y.Zhao is with Kent State University, USA (email: zhao@cs.kent.edu).

## I. Introduction

Visual Analytics is "the science of analytical reasoning facilitated by interactive visual interfaces" [67] (p.4), which focuses on developing human-computer methods and procedures for data analysis, knowledge building, and problem solving [40]. It is an applied research discipline that aims at creating methods of practical utility for different application domains, one of which is transportation. Obviously, the best results can be achieved when visual analytics researchers, who typically lack domain expertise, work in close contact with domain specialists. Unfortunately, such work has been limited in the transportation domain [26][27], even though visual analytics researchers have intensively worked with transportation-relevant data and developed a variety of methods and tools that could be useful for transportation domain researchers and practitioners. The consequences of the insufficient communication are two-fold. On the one hand, visual analytics researchers have only limited understanding of the problems, needs, and constraints of the transportation domain, which may decrease the potential utility and usability of the methods they develop. On the other hand, the transportation community has quite limited awareness of what visual analytics can offer. The ambition of this paper is to start building a bridge between the communities. We want to introduce visual analytics to transportation researchers and to present the spectrum of visual analytics works that we consider as potentially interesting to this audience. Hence, this is a survey of selected works in the visual analytics research field that deal with transportation-related data and tasks.

Before starting with the survey, it is important to explain the essence of visual analytics approaches and the conditions when they may be necessary. Visual analytics methods and procedures are designed for synergistic work between humans and computers where each side effectively employs its intrinsic capabilities. Specifically, humans employ their unique capabilities for creative thinking, making associations, and generating insights while computers process, aggregate, and mine data that would be too large for a human to effectively tackle alone. Interactive visual interfaces play a key role in these human-computer approaches, and visual representations are often the most effective way of conveying information to the human mind. By coupling these visual representations with interactions, users are enabled to explore information from different perspectives and levels of abstraction, thus associating distinct information pieces, and developing insights as information is perceived and interpreted.

A need for visual analytics approaches arises in situations that can be categorized as (1) new problems or (2) new opportunities. “New problems” can be subdivided into two sub-categories. The first sub-category includes problems for which no algorithmic solutions exist (yet). Here, the term “algorithmic” denotes not only computer-oriented algorithms but also established workflows with well-defined steps. The second sub-category includes problems for which some algorithmic solutions exist but have become ineffective or unsatisfactory because the problems have changed. A human-computer approach to a problem is necessary when the problem is insufficiently understood and/or ill-defined, and when it is not immediately clear how to tackle it. This calls for humans to engage in creative thinking, insight generation, and knowledge building.

“New opportunities” include the emergence of new types and sources of data or new technologies that may or may not be useful for solving existing problems in better ways. It is necessary to explore these new opportunities and find possible ways to benefit from them. This exploratory work is, obviously, a job for humans, who need appropriate support from computers. Transportation research is a well-established discipline in which numerous algorithmic solutions of transportation problems have been developed. However, the ongoing development of mobile devices, low cost sensors, driverless cars, as well as others, has led to information overload in the transportation sector. This data deluge presents transportation research with both new problems and new opportunities that call for human-computer approaches and makes visual analytics potentially helpful.

Here, transportation systems are seeing traditional problems transform due to substantial changes in the population structure (e.g., aging), spatial distribution (e.g., urbanization, urban sprawl, migration), people’s habits and lifestyles, and others. There is a need for gaining better understanding of the new or changed problems, which is leading to new opportunities arising due to the availability of large amounts of data that did not exist or were scarce in the past [16]. Historically, transportation analysis has relied on aggregate data measures, such as volume and speed data, available at the road segment level. However, the ubiquity of GPS enabled devices and data sharing has led to the creation of large corpuses of data related to movement. This includes not only data describing the movement of people (measurements from traffic sensors, tracks of vehicles, records of smart card transactions in public transport, etc.) but also data referring to population mobility, activities, and lifestyles (such as records of mobile phone uses and georeferenced posts in social media) [17]. The potential of these data types and sources for solving transportation problems needs to be explored.

Past surveys on visual analytics of movement data have focused heavily on the properties of trajectories and their underlying challenges [2, 11], other surveys have focused on collecting common visualizations used in traffic analysis (http://vizguide.camsys.com/), and a recent survey on traffic visualization [23] discussed visualizations to support traffic management and route planning. This survey aims to cover core concepts at the intersection of movement and transportation, filling a gap between the aforementioned surveys. The aim of this paper is not to cover all of the visualization and visual analytics methods being employed for transportation analysis; instead, our goal is to highlight recent work that is being done to support the next generation of datasets that can be used for intelligent transportation systems.

To this end, we have surveyed literature primarily in the visual analytics community in order to capture the current trends and research directions. Key venues surveyed include the IEEE VIS conference, the EuroVis conference, IEEE Transactions on Visualization and Computer Graphics, and Information Visualization. Papers dealing with trajectory and movement (which will be used interchangeably in this paper) were extracted. For completeness, we have also surveyed the last five years of papers from IEEE Transactions on Intelligent Transportation Systems (ITS) looking for keywords of visualization or visual analytics in the abstract. Several relevant papers on trajectory analysis from ITS will be discussed [31][33][41][46], in particular, those that have explored the integration of visualization and analytics techniques as part of the transportation analysis process [23][52][54]. We have found that relatively few works in the visual analytics community address specific transportation problems, and much of the transportation-related visual analytics research has been developed separately from transportation domain specialists. Thus, this survey aims to fill this gap in the literature and bring attention to the need for these communities to interact.

We divide the relevant works into four categories and present them in the following four sections.

Section II “Data” presents a typology of movement data that inventories data properties and possible issues, and describes data transformations relevant to analysis. Data issues were elicited from the experiences of the visual analytics researchers with numerous examples from the visual analysis of movement data, particularly related to transportation.

Section III “Movement and Transportation Infrastructure” discusses visual analytics approaches to analyzing *movement* data, specifically with respect to the movements of vehicles and pedestrians along transportation routes and movements of passengers within transportation systems (i.e., within transportation infrastructure). The data are considered from different perspectives and scales for exploring diverse aspects and features of movement behavior related to infrastructure in space and time.

Section IV “Movement and Behavior” refers to data concerning *people* who use or can potentially use transportation systems. Apart from data characterizing the use of transport by people, we also include data that do not refer to transportation directly but instead characterizes people’s general mobility behaviors, activities, and interests, which may be useful to take into account during transportation analysis and planning.

Section V “Modeling and planning” presents visual analytics works that go beyond the exploration and analysis of existing data to traffic modeling, forecasting, and planning.

After presenting the state of the art, we discuss, in section VI, the further tasks and directions for the visual analytics research for intelligent transportation systems.

The illustrations throughout the paper have been produced using the same example dataset consisting of GPS tracks of cars in Milan (Italy) collected during one week in April 2007, which were kindly provided for educational and research uses by company Octo Telematics (www.octotelematics.com). By this running example, we want to demonstrate the variety of possible ways to analyze the same data for deriving various kinds of knowledge. Please note that we cannot include illustrations for all techniques discussed in the survey.

## II. Data

Transportation research deals with a variety of data ranging from speed and volume counts on road segments to resiliency and capacity measures on infrastructure. Given that such data types and sources are well established in the transportation community, the data focus of this paper will be emerging data volumes which are primarily due to the ubiquity of GPS related devices, cell phones, smart cards, and other technologies that make it possible to acquire data reflecting movements of individuals. Such data can be used to give detailed insights into the movement of persons within their built environment and provide insight into their use of transportation infrastructure.

### *A. Data typology*

There are three fundamental types of spatio-temporal data [2]: spatial event data, trajectories of moving objects, and spatial time series. *Spatial events* are entities that emerge at certain spatial locations and exist for a limited time, such as a traffic jam or an accident. Some events, like traffic jams, may extend over large areas, which change over time. *Spatial event data* describe the spatial positions and extents, existence times, and thematic attributes of spatial events. *Trajectories* are chronologically ordered sequences of records describing the spatial positions of moving objects at different times, such as the moving paths of taxis, buses, or fleets. Additionally, the records may include values of thematic attributes that change as the objects move. *Spatially referenced time series*, or, shortly, *spatial time series* are chronologically ordered sequences of values of time-variant thematic attributes associated with *fixed* spatial locations or *stationary* spatial objects, such as segments of streets or public transport stops. For example, the time varying values of transportation volume or speed on a street segment generate spatial time series data.

Of the three data types, trajectories are among the most complex data in movement analysis. Trajectories describe positions of moving objects at sampled time moments. When the temporal and spatial gaps between these moments are small enough, the intermediate positions of the objects can be plausibly estimated by means of interpolation and/or map matching. Such data can be called *quasi-continuous*. Trajectories where recorded positions are separated by large time gaps, such that the intermediate positions cannot be reliably reconstructed, are called *episodic*. Quasi-continuous and episodic trajectories require different approaches for analysis [2]. An extreme case of episodic trajectories is data describing only trip starts and ends but not intermediate positions. Such data are usually referred to as *origin-destination (OD) data*, and well-known examples include data describing migration patterns or worker commutes.

While trajectories provide information on the movements of individual objects, aggregated trajectory data are spatial time series describing how many moving objects were present in different spatial locations and/or how many objects moved from one location to another during different time intervals. The time series may also include aggregate characteristics of the movement, such as the average speed and travel time. Time series describing the presence of objects are associated with fixed locations, and time series describing aggregated moves, often called *fluxes* or *flows*, are associated with pairs of fixed locations.

### *B. Examples of visual representation of different data types*

The spatial aspect of the different types of spatio-temporal data can be represented visually on maps. Spatial events are often represented by dot symbols drawn at the event locations (Fig. 1A), when the spatial extents and shapes are negligible, irrelevant for analysis, or unknown; otherwise, events can be represented by polygons. Trajectories are typically represented by lines connecting the object positions (Fig. 1B).

A visualization method called the space-time cube (STC) can simultaneously represent the spatial and temporal aspects of spatial events and trajectories (Fig. 1C). Two dimensions of the STC represent the geographic space, and one dimension represents the time. The base of the STC usually contains a map providing the spatial reference. In Fig. 1C, time is represented by the vertical dimension of the STC. The time axis is directed from bottom to top. Spatial events and points of trajectories are placed in the STC according to their spatial positions and times. The points of trajectories are connected by line segments in chronological order. A space-time cube display requires interaction, allowing rotation of the scene as well as panning and zooming to adjust the viewpoint. These interactions are used to improve the perception of spatio-temporal patterns. Still, the STC display often suffers from visual clutter and over-plotting of visual symbols. To be effective, the STC is often used in combination with interactive filtering and clustering applied to events or trajectories as a means of clutter reduction and aggregation.

For spatial time series, there is no convenient visualization method to represent both the spatial and temporal aspects. A map can show the spatial distribution of the presence of moving objects and/or their flows between locations corresponding to one time step (interval). Multiple time intervals need to be represented by a sequence of maps. When the time series are short, the maps can be put side by side; otherwise, map animation is used.

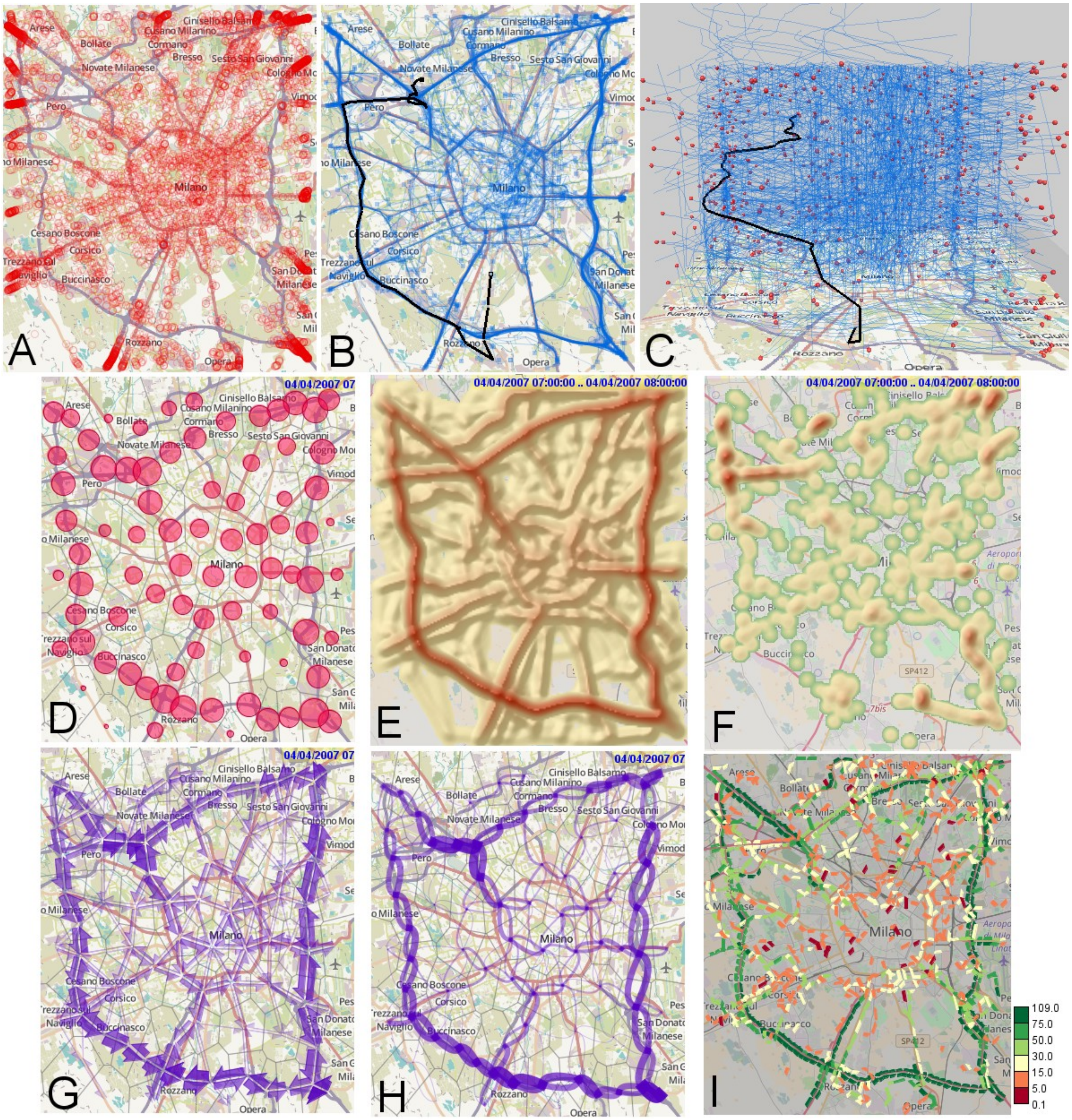


Fig. 1. The types of spatio-temporal data and their typical visual representations. A: Spatial events (e.g., car stops) represented on a map by dot symbols. B: Trajectories of cars represented on a map by lines. One selected example trajectory is marked in black. C: Spatial events and trajectories represented in a space-time cube by dot symbols and lines, respectively. The same trajectory as in image B is marked in black. D: Counts of moving objects (cars) in different spatial compartments in one time interval are represented by proportional circle sizes. E: A continuous density map represents the distribution of the traffic density across a territory. F: A continuous density map represents the distribution of spatial events (slow movement of cars). G: Flows of cars between spatial compartments in one time interval are represented by half-arrow symbols with the widths proportional to the flow magnitudes. H: The same flows are represented by curved lines; the curvature is higher at the end (destination) of a flow. I: Average speeds of cars are represented by color coding.

For one time step of a time series, the spatial distribution of the presence of moving objects can be represented on a map by symbols or diagrams positioned at different locations over the territory or in different territory compartments (Fig. 1D) with the sizes proportional to the counts of the objects or other characteristics of the presence, such as the average duration of staying. Another possible representation of object presence is continuous density map (Fig. 1E). In such a map, the variation of colors or shades encodes the variation of presence or movements across a territory. A density map effectively

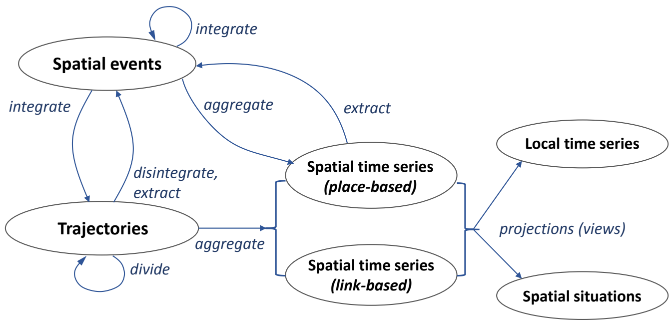


Fig. 2. Possible transformations between the types of spatio-temporal data.

reveals existing traffic channels and their relative importance but does not show the movement directions. An animated density map can represent the variation of the movement density over time. Each image in the animation shows the density in one time interval. Continuous density maps can also be used to represent the distribution of spatial events over a territory in a chosen time interval. For example, Figure 1F shows the distribution of slow movement events in time interval 07:00-08:00 on April 4, 2007. The hot spots seen in the map may correspond to traffic congestions.

Flows between locations or territory compartments are typically represented by linear flow symbols connecting the locations or compartments. The flow directionality is signified by arrows at the line ends (Fig. 1G) or by variation of the line curvature (Fig. 1H). The widths of the lines are proportional to the flow magnitudes, that is, the counts of the objects that moved, or the volumes of transported goods, or other numeric characteristics of the flows. Such maps are commonly called *flow maps* [43][68]. Animated maps of presence and flows are often combined with temporal displays, such as a time graph, showing the variation of the presence/flow magnitudes for different locations or links between locations, respectively.

Besides flow magnitudes, other aggregate characteristics of movements can be represented by varying the appearance of flow symbols. For example, average speeds are represented in Fig. 1I by color coding using a diverging color scale from red to green for speed values from low to high.

### *C. Data transformations*

The different types of spatio-temporal data do not exist in isolation. There are techniques for transforming one data type to another [2]. Data transformations may be needed to prepare data for analysis methods and/or to align the spatio-temporal phenomenon reflected in the data at varying scales.

A summary of possible transformations between the spatio-temporal data types is presented in Fig. 2. The left part of the diagram shows the tight relationships between spatial events and trajectories. In fact, trajectories consist of spatial events: each record in a trajectory of an object represents a spatial event of the presence of this object at a specific location at some moment in time. Trajectories are obtained by integrating spatial event data: for each object, all its position records are linked in a chronological sequence. Reciprocally, trajectories can be transformed to spatial events either by full disintegration back into the constituent events or by extraction of particular events of interest ([2], sections 3.5, 5.2), such as stops, sharp turns, or encounters of two or more objects.

**Spatial Events:** Multiple spatial events that are close in space and time can be united into more complex spatial events. For example, a spatio-temporal concentration of many vehicles reducing their speed during a small time window may be treated as a single event of traffic congestion. Such composite spatial events can be detected and extracted by means of density-based clustering ([2], section 6.1). To represent a composite event as a single entity, a spatio-temporal envelope may be built around the constituent events [8].

**Trajectories:** Often, trajectories of moving objects are available as unitary sequences of recorded positions extending throughout the whole period of observation, including the time intervals when the objects did not move. For certain analysis tasks, it may be reasonable to separate movements from stops and divide full trajectories into smaller trajectories that represent the movements (trips) between the stops. There may also be other reasons and criteria for dividing trajectories ([2], section 3.2).

**Spatial Time Series (Place-Based):** Spatial time series can be obtained from spatial events or trajectories through spatio-temporal aggregation. For discrete spatial aggregation, the underlying regions in which the events or trajectories take place can be divided into compartments, and time is divided into intervals. For each compartment and time interval, the spatial events or moving objects that appeared in the compartment during the associated time interval are binned together and counted. Other aggregate statistics can also be computed. The result is a place-based time series in which temporal sequences of aggregate values are associated with the places (i.e., spatial compartments). From such spatial time series, in turn, it is possible to extract spatial events ([2], section 7.2.5), for example, events of high traffic density or events of extremely low average speed.

**Spatial Time Series (Link-Based):** Trajectories can also be aggregated into link-based time series: for each pair of compartments and time interval, the objects that moved from the first to the second compartment during this time interval are counted and aggregate characteristics of their movements (e.g., the average speed) are calculated.

**Local Time Series and Spatial Situations:** Discrete place-based and link-based spatial time series can be viewed in two complementary ways. On the one hand, they consist of temporally ordered sequences of values associated with individual places or links, i.e., local time series. On the other hand, a spatial time series is a temporally ordered sequence of the distribution of spatial events, moving objects, or collective moves (flows) of moving objects over the whole territory and the spatial variation of various aggregate characteristics. These distributions are called “spatial situations” [2].

**Spatial situations represented as continuous fields:** Continuous spatial aggregation (as in Fig. 1E, F) is done using a raster, i.e., a regular grid dividing the territory into small cells. As in discrete aggregation, counts or other aggregates are obtained for the cells. Then, spatial smoothing is applied, which combines the value in each cell with the values in the surrounding cells using a special weighting function (kernel function). The function defines the manner in which the weights of the surrounding cells decrease as the distance to the central cell increases. The result is a smooth density field. Continuous spatial aggregation can be combined with discrete temporal aggregation based on time division into intervals. A density field is generated for each time interval and represents the distribution of spatial events or movements during that interval. Hence, the result of this aggregation is a time series of spatial situations. Unlike the case of discrete spatial aggregation, such spatial time series cannot be viewed as a set of local time series.

**Other transformations:** Apart from these standard transformations between or within the different types of spatio-temporal data, it is possible to transform data to a completely different representation, which may be beneficial for particular tasks. For example, Chu et al. [24] transform trajectories of taxis into sequences of the names of the traversed streets and apply text mining methods for discovery of "taxi topics", i.e., combinations of streets that have a high probability of co-occurrence in one taxi trip. The extraction of "taxi topics" is done for different time intervals. By investigating the temporal evolution of the topics, it is possible to understand where people travel in different times of the day and days of the week. Al-Dohuki et al. [1] transform taxi trajectories into texts consisting of street names and text labels denoting taxi speeds (low, medium, and high). This representation is used for supporting queries to a trajectory database where users can formulate queries by specifying street names and/or speed characteristics. The queries are performed by means of a text search engine. Furthermore, a discrete representation of aggregated movements of flows between places can be treated as a graph, to which graph analysis methods can be applied [32][49].

As such, these various transformations enable the comprehensive analysis of movement data from multiple complementary perspectives [11].

### D. Data enrichment and integration

Spatial event data and position records in trajectories can be enriched by computing a number of derived attributes [8]. Thus, for quasi-continuous movement data, it is possible to derive movement speed, acceleration, direction, and turn from the available positions and times, as well as time interval-based measures, such as the path length and displacement by time intervals of a given length. For events and trajectory positions, it is possible to compute attributes characterizing their neighborhood, such as the number of other events or moving objects in the vicinity, specified by given spatial and temporal distance thresholds, and the distances to the nearest neighbor.

Movement and event data can also be enriched through integration with other spatial, temporal, or spatio-temporal data based on commonality or proximity of the spatial and/or temporal references in the different datasets. An example is attaching weather attributes to positions of vessels [53].

### E. Exploration of data properties and quality issues

To assess the suitability of data for analysis, it is necessary to investigate the data quality, attributes, and distribution. Data quality issues, structure and feature relationships can often be revealed by appropriate visualizations ([2], section 9.2). In spatiotemporal data this may refer to misaligned temporal resolution, temporal regularity or irregularity, presence of temporal gaps, varying spatial resolutions and the presence of spatial gaps, issues concerning identities of moving objects, properties related to the method of data collection, positioning errors, and others. A typology of possible quality problems that can be encountered in movement data is introduced by Andrienko et al. [6], which also demonstrates how visualizations can reveal such problems.

As such, data being studied has underlying uncertainty that should be conveyed to the domain experts. Furthermore, much of the trajectory data being captured is from taxis and trucks as opposed to regular passenger cars, which may bias the data and add more uncertainty. Visualizing uncertainty has been listed as an ongoing challenge in visualization [39], and a recent survey [42] discusses methods of uncertainty visualization specifically in the context of spatiotemporal data. While outside the scope of this survey, such techniques should also be considered when designing for future transportation analytics systems.

### F. Dealing with large data volumes

Currently, data being collected by GPS enabled devices are characterized by large volumes and such data volumes pose serious challenges to visual analytics methods and software tools. To enable interactive querying and analysis, data need to be quickly accessed, extracted, transformed, and visualized (ideally at interactive rates, ~10ms). This requires an effective data management system. Given that much of the GPS data being captured is in the form of trajectories, existing systems for transportation data analysis do not always provide the required infrastructure, which has lead visual analytics researchers to develop tailored approaches including specialized data indexes [26][52] and hash structures [70].

Apart from effective data management, visual and interactive techniques and analysis methods need to be appropriately designed for dealing with very large amounts of data. Data aggregation is a common technique, in particular, adaptive aggregation depending on the spatial and/or temporal scale of the current view. Initially, large amounts of data are visually presented in an aggregated way for an overview. As the user zooms in and focuses on particular areas and/or time periods, more details are shown [26][52]. A related problem is to reduce display clutter when many moving objects need to be shown. This can be solved by grouping (clustering) spatially close objects and showing aggregated data for the clusters [62]. A strategy to extend the capacity of analysis methods such as clustering beyond the limitations of computer RAM is to perform an initial analysis on a subset of the data and use the results to interactively build a model (such as a classifier) that can be automatically applied to the remaining data [9].

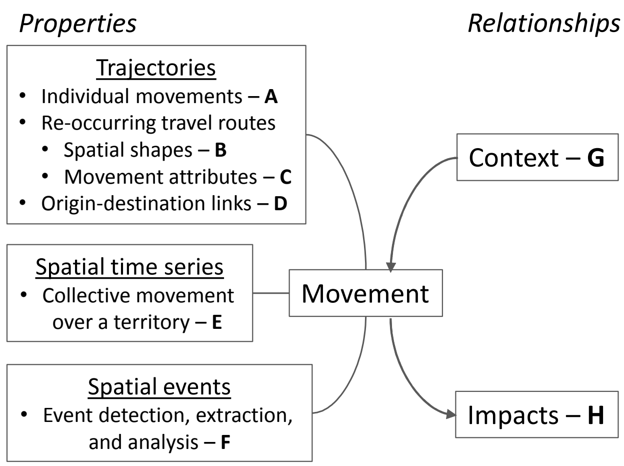


Fig. 3. Schematic representation of the structure of Section III.

Other applications require dynamic processing, analysis, and visualization of real time streaming data. This necessitates the development of methods for incremental analysis and visualization, in which previous analysis results and visualizations are continuously updated using new data. An example is the real-time detection of complex events, such as traffic jams, composed of multiple elementary events [13]. An incremental algorithm for clustering spatial events detects spatio-temporal concentrations (clusters) of events in real time and tracks the evolution of the clusters. A dynamic visual display updates to show the current states of the clusters and their continuing evolution.

## III. Movement and Transportation Infrastructure

In this section, we discuss visual analytics research on analyzing movements of vehicles and pedestrians within transportation infrastructures. The structure of the section is schematically represented in Fig. 3. Movement data can be represented in the form of trajectories, spatial time series, and spatial events, each representation being suitable for studying different aspects of movement [2]. Subsections A-D focus on the representation of movement by trajectories, which enables exploration of individual movements (A), travel routes in terms of spatial shapes (B) and dynamic movement attributes (C), as well as links between trip origins and destinations (D). Subsection E focuses on spatial time series representing properties of collective movement over a territory, and subsection F deals with events pertaining to movement. While subsections A-F refer to various aspects and properties of movement per se, subsections G and H refer to relationships between movement and other phenomena and entities. Subsection G deals with analyzing external factors (context) influencing movements, and subsection H considers the works on analyzing negative impacts and risks associated with vehicle movements.

### A. *Details of individual movements*

Here we present the visual and interactive techniques designed for a detailed exploration of movements, usually at a small spatial scale. The techniques enable the analyst to see the movements and characteristics of individual objects and select particular objects for a close inspection.

TripVista [30] represents individual movements of vehicle and pedestrians at a road intersection by polylines colored according to the types of the moving objects or the movement speed. By interacting with the display, the user can select trajectories with particular shapes. Pu et al. [56] represent movement characteristics of individual vehicles by specially designed glyphs.

Unique interactive techniques of FromDaDy [35] are applied to a large number of individual aircraft trajectories for flexible selection and extraction of subsets and parts of trajectories for separate exploration. An interesting feature of the system is the representation of 3D trajectories in 2D projection views. This enables the interactive selection of trajectories based on altitude or the speed of the accent/descent.

### B. *Variety of taken routes*

Techniques in this subsection focus on the visual representation of travel routes of moving objects in geographic space. Interactive techniques and clustering allow the analyst to assess the diversity and repeatability of the routes, find frequently taken routes, reveal the possible ways for getting from an origin to a destination, and explore the differences between alternative routes connecting the same locations.

With TrajectoryLenses [45], routes are explored in a purely visual and interactive way. The user can select and see the trajectories going from a selected area of origin or coming to a selected destination area, or all routes from a selected origin to a selected destination. More sophisticated queries specifying intermediate waypoints are also possible. Interactive selections are also enabled in a system designed by Liu et al. [51], and work by Liu et al. also visualizes aggregated information about route diversity over the entire territory under analysis. The aggregation groups elementary locations into larger areas, for which incoming and outgoing diversity scores (i.e., the numbers of distinct routes followed by the incoming and outgoing the trajectories) are computed and visually represented. The user can select an origin-destination pair and investigate the respective trajectories using a detailed view. It is also possible to select a road segment and explore the diversity of the routes going through this segment.

Rinzivillo et al. [58] use density-based clustering for grouping trajectories according to the closeness of their origins and/or destinations or according to the similarity of the routes they follow. Particularly, clustering according to route similarity finds frequently re-occurring routes, which can be visualized using aggregate flow symbols (Fig. 4). Later work by Andrienko et al. [9] proposes a scalable variant of the method, in which clustering is applied to a subset of trajectories loaded in RAM. Based on the clustering results, a classifier for identifying the cluster membership of an arbitrary trajectory is interactively built. It is then applied to the whole set of trajectories stored in a database.

Zheng et al. [79] propose a set of techniques supporting analysis of routes of passengers in a public transportation system. This includes specific computations, such as the travel efficiency of a route, which accounts for the riding, waiting, and transfer times. A tree-like visualization, called isotime flow map, shows efficient journeys (by travel time) starting from a selected area. A map-based isochrone view shows, for a selected origin, the reachability regions corresponding to a given time budget.

As mentioned earlier (section II.C), trajectories can be transformed to a text-based representation so that a trip is represented as a sequence of street names. Then, the routes followed in the trips can be analyzed using techniques for text analysis, such as topic modeling [24], and visualized using text-oriented visual displays, such as text cloud [1][24]. A "topic", which consists of names of streets that frequently co-occur in one trip, may evolve over time. Topic evolution is reconstructed by means of computational techniques that match topics extracted from different time periods. The evolution is visually represented on temporal displays.

*C. Movement dynamics along a route*

Here we focus on techniques designed for analyzing dynamic attributes of movements (speed etc.) along a particular travel route or channel, which is typically considered to be a line segment, for example, along a street, ship lane, or metro line. The analyst can see and explore the variation of the attributes over time and across multiple trips.

To show the variation of movement characteristics, such as speed or tortuosity, within multiple trajectories following similar routes or going through the same street, Tominski et al. [69] designed a 3D view (Fig. 5) in which the trajectories are put on a base map in a stack. Each trajectory is represented by a colored ribbon where colors encode attribute values. Additionally, the variation of the attribute values over time in the entire trajectories or at a selected position is represented on a circular display (Fig. 5, bottom right). Case studies focused on the detection of traffic congestion on streets and anomalies in vessel traffic. A similar 3D representation is used by Itoh et al. [36] to show the variation of passenger flows along the lines of a metro network. The ribbon widths are proportional to the numbers of the passengers, and colors represent the level of crowdedness. In addition to the map-based 3D view, there is a tabular temporal display with the rows corresponding to the metro lines and columns to time

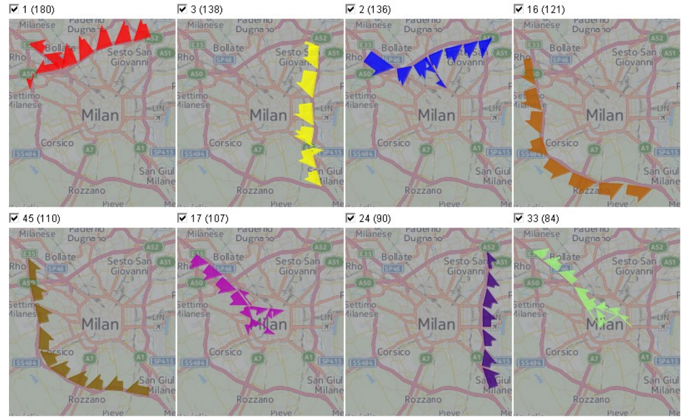


Fig. 4. Major routes taken by cars in a city have been revealed through density-based clustering of trajectories according to similarity of the routes. The clusters are represented in a summarized form using flow symbols.[58]

intervals. The variation of passenger flow characteristics is represented by color coding. Similar encoding is applied in Trips Explorer and Stops Explorer [55] for visualization of the public transport performance along a selected route. In the displays with two dimensions representing the time and the sequence of stations, color variation is used to show various characteristics such as trip frequency, waiting times, speed of the movement, deviations from the schedule, delays, etc. To reduce visual clutter, the displays are smoothed by means of kernel density estimation techniques.

Wang et al. [70] provide a map-based interface for selecting subsets of trips going through a street segment. Movement characteristics in the selected trips are shown in separate displays, such as scatter plots and histograms. Wörner and Ertl [77] show the dynamics of speed or other attributes on a graph where the horizontal axis represents the route or street length and the vertical dimension represents the attribute values. Apart from lines corresponding to different trips, a line connecting the mean values and a standard deviation envelope are shown.

Qiang et al. [57] propose an original technique for simultaneously representing movement characteristics in full detail and at different levels of aggregation. In a 2D display, the horizontal dimension represents time or the street extent, and the vertical dimension corresponds to different levels of aggregation, from maximal detail at the bottom to maximal aggregation (i.e., a single value) at the top. The display appears as a continuously colored triangle where colors encode attribute values at different levels of aggregation.

Sun et al. [66] show the weekly variation of traffic amounts on street segments directly on a map by drawing time series graphs along the segments. Traffic flow magnitudes in two opposite directions are shown on two sides of the time axis and in two distinct colors. For journeys by public transport, Zeng et al. [79] show the travel times by segments of alternative routes connecting a selected pair of origin and destination locations. The routes are shown in a tree-like display where the horizontal dimension represents the cumulative travel time and the tree branches represent different routes. The variation of the travel times over a day is shown on circular diagrams.

### D. *Linking origins to destinations*

This subsection focuses on methods for supporting the analysis of origin-destination (OD) travel data, i.e., data specifying the locations and times of trip starts and ends. The full trajectories are either not available or not relevant to the analysis. OD data are often aggregated into matrices or flows, such that each matrix cell or each flow represents all trips from some origin to some destination. Both detailed and aggregated OD data pose a great challenge to visualization. Matrix views may be insufficient for analysis as they do not convey spatial patterns. On a map, it is very hard to represent multiple intersecting moves across a territory in a legible and easily understandable way. This problem pertains also to episodic trajectories that have some intermediate points between origins and destinations. As these points are separated by large temporal and spatial gaps, each segment of such a trajectory needs to be treated in the same way as an OD move. Researchers apply clustering techniques to simplify OD flow maps [68] or invent alternative techniques for representing connections between origins and destinations [75].

Spatial simplification can be achieved by grouping the origin and destination locations into larger regions and aggregating the trips into flows between the regions. Regions can be defined by means of spatial clustering of neighboring locations [29], possibly, taking into account the strengths of the flows between them [49]. Flow data can also be simplified by grouping and aggregating spatially close OD flows using hierarchical clustering [81]. Another approach is visual simplification by edge bundling (e.g., [25]), i.e., merging of spatially close flows and representing them by branching lines. On a geographic map, this works well only for showing flows from one or two locations or in special cases, e.g., when radial flows from/to one location prevail over all others, as the flights between Paris and other cities in France [25].

To represent time series of flow variations while reducing map clutter, Boyandin et al. [20] propose a visualization consisting of two maps and a table display with the rows showing time series of flow magnitudes. The rows are connected by lines with the flow origins in one map and destinations in the other map. This technique is suitable for tracing individual links and viewing their local time series, but it does not show the spatial patterns of the flows.

To avoid showing flows by intersecting lines, OD maps have been proposed [75]. They are based on space transformation in which the locations are arranged in a matrix so as to minimize the distortions of their relative spatial positions with respect to each other. Each location is represented by a matrix cell, which is filled with a small matrix of the same structure as the overall matrix. The inner matrix represents the flows from/to this location to/from all other locations. Such display is free from occlusion, but the space distortion complicates the perception, and the overall spatial pattern of flows is broken into multiple location-specific patterns. Recently, it has been proposed to aggregate OD data in a way that not only reduces the data dimensionality for efficient interactive analysis but also enables visual representation by means of diagrams rather than intersecting

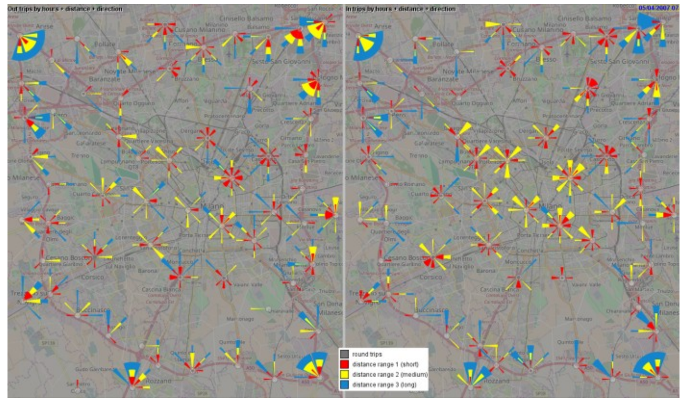

Fig. 6. Aggregated outgoing (left) and incoming (right) car trips to/from different directions and distance ranges are represented by diagrams with segment widths proportional to the flow magnitudes.[7]

flow lines [7]. The diagrams are positioned at the places of trip origins (Fig. 6, left) or destinations (Fig.6, right) and show the counts of trips to/from different directions and distance ranges. The temporal variation of the trip distribution is studied using temporal clustering of spatial situations.

Rather than trying to present OD trips over the whole territory in a synoptic way, Ferreira et al. [26] focus on supporting interactive queries to a database of OD data (taxi trips). The user can specify a time interval, origin and/or destination regions, or trip direction. The system selects the trips satisfying the query and shows statistics of their characteristics on graphical displays. The origins and destinations of the trips are represented on a map by dots of two distinct colors. Jiang et al. [38] represent the spatial distributions of the trip origins and destinations by density maps. For a user-selected region, characteristics of the incoming and/or outgoing trips are visually represented on multiple graphical displays.

### E. *Collective movement over a territory*

This subsection presents approaches to support an overall view of the movement distribution and properties over a large territory based on aggregation of individual movements. Different methods of spatial aggregation produce continuous fields of movement density or discrete representations of the presence of moving objects by space compartments and collective movements (flows) between the compartments. The aggregation is also applied to subsets of data, which can be selected by interactively setting spatial, temporal, and/or attribute constraints.

To support an overall view of movement over a territory, information from multiple trajectories needs to be aggregated over space. As mentioned in section II.C, there are two approaches to spatial aggregation, continuous and discrete. In continuous aggregation, a smooth density surface is generated using kernel density estimation techniques. For aggregation of trajectories, speed variation is integrated in the kernel convolution along the path [73], and the kernel width is automatically adapted according to zooming and panning operations [47]. On top of a density map, animated particles can represent the movement directions [59]. The user can

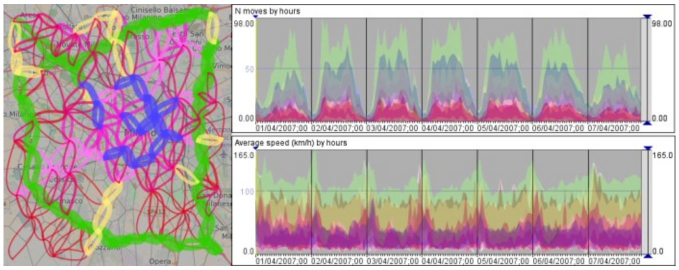


Fig. 7. Flows between spatial compartments have been clustered according to the similarity of the local time series of the flow magnitudes and speeds. Left: the flows on a map are colored according to the cluster membership. Right: the temporal variations of the flow magnitudes (top) and mean speeds (bottom) by the clusters are represented on time graphs.[2]

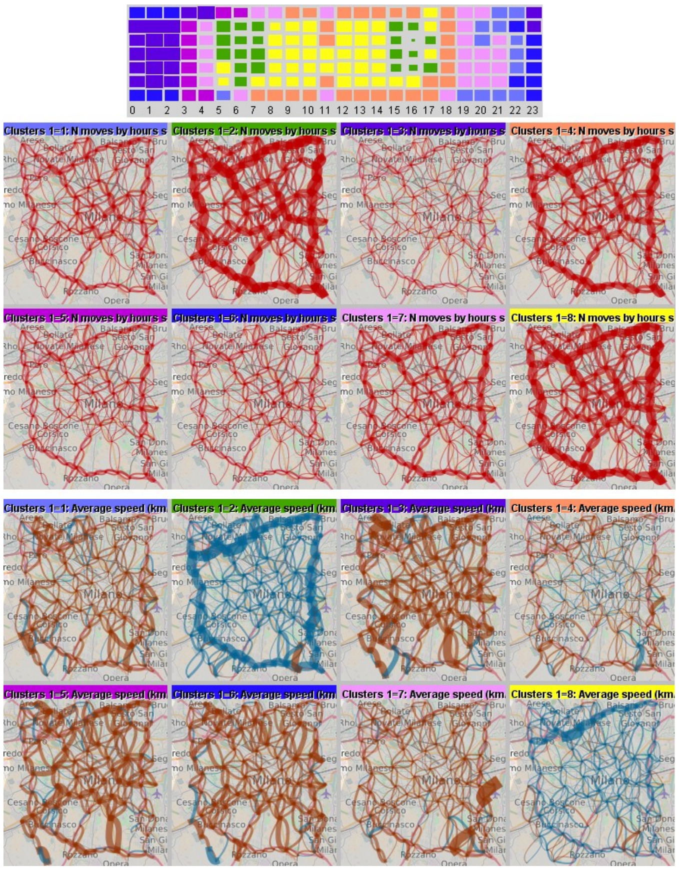


Fig. 8. Hourly time intervals over a week have been clustered by the similarity of the spatial situations in terms of the flow magnitudes and average speeds. In a time matrix at the top, the rows correspond to the days from Sunday to Saturday and columns to the day hours. The time intervals are represented by rectangles colored according to the cluster membership; the sizes show the closeness to the cluster centers. Below, representative spatial situations for the clusters are shown by flow maps. In the upper set of 8 maps, the widths of the flow symbols are proportional to the mean flow magnitudes. The lower set of 8 maps represents how the mean speeds in the clusters differ from the median mean speed attained on the links. Positive and negative differences are encoded by proportional widths of flow symbols colored in brown and blue, respectively. [2]

interactively select particular flows for viewing and comparing their variations over time, which are represented on linear and circular histograms [59]. Comparison of two density maps, e.g., corresponding to different time intervals or different types of moving objects, can be supported by subtracting one map from another and encoding positive and negative differences by shades of two color hues [48].

The concept of density maps can be extended to representing not only densities but also other attributes, such as traffic velocities [63] or the number of taxi customers [52]. Several density images built with different parameter settings or representing different attributes can be combined in a single composite density map using special operators [63]. Examples show that such a map can effectively differentiate moving and anchoring vessels or highlight anomalous movements.

Discrete aggregation, as explained in section II.C, produces place-based and link-based spatial time series, which can be viewed and analyzed in two complementary ways: as a set of spatially distributed local time series and as a chronological sequence of spatial situations. In order to provide a comprehensive understanding of the movement behavior over space and time, both views may need to be considered. Andrienko et al. [2][5] propose an approach involving two-way clustering, where a partition-based clustering algorithm is applied to the local time series and the spatial situations. The application to the local time series results in clusters of places or links characterized by similar value variations (Fig. 7). The application to the spatial situations results in clusters of time intervals characterized by similar spatial situations (Fig. 8). In this way, spatial and temporal simplification and abstraction are achieved, which facilitates comprehending the overall behavior over space and time.

Local place-based time series can be visually represented by diagrams drawn on top of a background map, for example, by circular diagrams representing the variation of movement characteristics over a time cycle. Diagrams may consist of concentric rings corresponding to different days while each ring represents the daily time cycle [52][56].

Wang et al. [72] spatially aggregate trajectories by traffic monitoring cells, which are distributed over the street network but do not cover the whole network. Each cell corresponds to a single movement direction. The result of the aggregation is treated as a graph with the nodes corresponding to the cells and the edges to the links between the cells. The cells are represented on a map by glyphs showing the movement directions, flow volumes, and speeds. The links are represented by lines with the widths proportional to the flow volumes. The temporal variation of traffic characteristics for selected cells or links can be explored using additional displays. Huang et al. [32] exploit a graph-based representation to an even greater extent. In Huang et al.'s work, street segments are represented by graph nodes where the links and their weights are defined based on the existing taxi trajectories. Calculation of graph centrality metrics, in particular, pagerank and betweenness, is applied to the street segments. The results, which are visualized on maps, characterize the time-varying importance of the street segments.

### F. Events

Movement includes many events, some of which may require special attention and analysis, in particular, negative events such as incidents, failures, dangerous movements, and congestions. Events requiring analysis may not be explicitly specified in data. There are interactive techniques for the extraction of events that need to be studied from movement data and methods for analyzing the temporal patterns and trends in the event occurrences over space.

Fredrikson et al. [27] described a system for the visual exploration of traffic incidents using spatial, temporal, and categorical (by incident type or other attributes) aggregation of data reflecting individual incidents. A web-based system with similar functionality was developed more recently [74]. In these works, the events were explicitly specified in the data. There may be a need to detect abnormal events by analyzing other kinds of data, such as trajectories of moving objects. T-Watcher [56] supports the visual detection of various anomalies in traffic using aggregated and detailed views, and work by Hamad and Quiroga [31] demonstrates the use of geographic information systems to explore transportation management performance measures in San Antonio Texas. Their focus was on performance evaluation of incident detection algorithms using spatial visualizations.

Furthermore, it may be necessary not only to detect specific events but also to extract them (i.e., separate from the remaining data) for further analysis. This can be done using interactive filtering techniques. A general procedure [2][8] consists of four steps: (1) compute relevant dynamic attributes; (2) define thresholds separating abnormal values from normal; (3) use these thresholds in constructing a filter, which may also be based on several attributes; (4) extract the points or segments of the trajectories that satisfy the filter. An example is the extraction of points with low speed values from vehicle trajectories for the detection and analysis of traffic jams [8][71].

Points or segments extracted from trajectories are elementary events representing particular states of individual moving objects, such as stop, slowing down, or approaching other objects. These elementary events may not be of interest per se, but they may be parts or indications of important complex events. For example, a spatio-temporal concentration (cluster) of vehicles decreasing in speed may signify a traffic jam. To identify the locations and spatio-temporal boundaries of such complex events, spatio-temporal density-based clustering can be utilized [2][8]. A special incremental event clustering algorithm capable of working in streaming settings for detecting event clusters in real time and tracing their further evolution has been proposed [13]. One of the use cases is the online detection and tracking of traffic jams. To represent a complex event as a single object, a spatio-temporal envelope (such as a convex hull) is built around the elementary events included in the complex event.

Wang et al. (2013) have developed specific techniques for analyzing traffic congestions. Taking into account the spatial connections between street segments and the times of traffic slowing down and assuming backward propagation of traffic jams (i.e., in the direction opposite to the movement direction), they build a jam propagation graph. The graph shows how an emergence of a traffic jam on a street segment affects other street segments over time.

### G. *Contextualizing movement*

Here we touch upon the visual analytics approaches for analyzing how movements are affected by external factors (context), such as weather or emergency events. The approaches involve joint analysis of movement data and data concerning the spatial and/or temporal context of the movement. Links to relevant contextual data are established based on the spatial and temporal references present in movement data.

Lundblad et al. [53] attach weather data to positions in vessel trajectories. The user may select some ships and see the weather attributes along their routes in a time graph. The user may also select a time moment and see the weather attributes for all ships in a parallel coordinates plot. Buchmüller et al. [21] have developed a system that allows users to explore the relationships between the directions of aircraft landings at an airport and the weather parameters to evaluate the noise impact of airplane landings on the surrounding areas. Users can choose time intervals of interest and see the aircraft trajectories and weather information. Furthermore, the system includes a model that predicts the expected distribution of the arrival directions for user-specified weather conditions.

Weather conditions, in particular, the direction and speed of the wind, not only determine the directions of aircraft takeoffs and landings but also affect the ground speeds of airplanes as they fly. The wind impact is clearly seen in a visualization of the aircraft ground speeds against the headings; moreover, wind parameters can be extracted from dynamic attributes of several airplanes flying over the same region in different directions [34].

For detecting and exploring the impacts of extraordinary events, such as disasters, accidents, and public gatherings, on the use of public transport (metro), Itoh et al. [36] visualize deviations from the average passenger flows on different metro lines by time intervals. Upon detecting an anomaly, the user can obtain related information from social media (Twitter). For user-specified time intervals and metro stations or lines, the system finds related tweets and shows the frequent keywords, which may explain the reasons for the anomaly.

### H. *Impacts and risks*

Unfortunately, transport systems bring not only various benefits but also numerous negative impacts on the environment, society, and economy. In addition to the issues pertinent to normal transportation activities, illegal activities and unruly behaviors pose further dangers.

The work of Buchmüller et al. [21] focuses on the problem of noise from aircraft landings at Zurich airport, which affects people living in Germany close to the Swiss border and causes an ongoing conflict between the German and Swiss sides. Buchmüller et al. developed a system for the visual exploration of aircraft landing data and, in particular, checking whether the pilots adhere to the existing rules, detecting rule violations, and examining the context (time and weather conditions) in which they occurred.

Scheepens et al. [60] focus on the problems of safety and security in maritime transport. They developed an interactive visual interface to an automated inference engine that detects dangerous or suspicious behaviors of vessels and raises alarms. The purpose is to present the rationale for the alarms in an easily perceivable and understandable way. An explanation graph shows the reasoning structure and the probabilities of different hypotheses according to the available evidence (observations). The observations are represented in a matrix showing also the confidence levels and agreement or disagreements between the observations. The matrix rows are connected to graph nodes showing which observations contribute to which hypotheses. Scheepens et al. also presents several use cases involving the detection of possible environmental hazards, reckless behavior of a vessel, and suspected smuggling.

## IV. MOVEMENT AND PEOPLE'S BEHAVIOR

While the previous section focused mostly on transportation means, this section focuses on people as actual or potential users of transportation means and services.

### A. *Use of transport*

This subsection considers visual analytics approaches to analyzing the use of transportation means by people. The existing techniques analyze the spatial and temporal patterns and trends, reveal behavioral differences between user groups, and relate the use of transport to the spatial and temporal context and people's activities.

Human mobility behaviors over public transit systems are commonly explored to identify commute patterns and reveal behavioral differences. For example, Wood et al. [76] visualize the dynamic patterns of a bicycle hire scheme in London. Flow maps with symbols provide overviews of bicycle traffic flow structures, and an origin-destination map is used to show details on demand. The status of docking stations over space and time is further visualized in a grid view, and patterns of the bicycle hire program revealed insight into how different populations use the bicycle hires. The spatio-temporal patterns of bicycle trips over a long time period were also investigated using aggregation of OD data by trip directions and distance ranges and clustering of spatial situations from different time intervals [7].

Recent work by Beecham and Wood [18] further explores the bicycle hire scheme to analyze gendered cycle behaviors with regard to spatial, temporal, and customer-related variables. They found that female customers' usage characteristics seem to be related to weekend usage and parks, where men appear to utilize bike hires for commuting. Other work in OD pairs has focused on Bluetooth data. Laharotte et al. [46] used Bluetooth detectors in Brisbane to create B-OD matrices to describe the dynamics of a subpopulation of vehicles to characterize urban networks.

Further exploration of customer behavior includes van der Hurk et al. [33] which presents a methodology for extracting passenger routes based on smart card data from the Netherlands Rail System. This work demonstrates how passenger service, based on passenger route choice, can be analyzed based on the route detection mechanism. In a similar direction, work by Kieu et al. [41] explored the use of smart card data for passenger segmentation. Here, the goal is to group passengers of similar travel patterns to identify market segments for transit authorities to help understand utility (or disutility) for improved services.

To further support the analysis of the use of transportation means, Kruger et al. [45] develop an interaction technique, TrajectoryLenses. Complex filter expressions are supported by the metaphor of an exploration lens, which can be placed on an interactive map to analyze geospatial regions for the number of trajectories, covered time, or vehicle performance. Case studies explored usage behavior of people that employ electric scooters for daily travelling. Another work by Krueger et al. [44] enriches the trajectories of the scooter users with semantic information concerning the visited places to infer users' activities and travel purposes. Semantic insights of points of interest are discovered from social media services. The uncertainties in time and space, which result from noisy, unprecise, and missing data, are visually analyzed by the geographic map view and a temporal view of OD patterns. In this way, people's activities can be related to nearby locations and semantically tied to the point of interest data.

Other work has focused on transforming the geographic coordinates of taxi trajectories into street names. In this way, the movement of each taxi becomes a document consisting of the traversed street names [24]. The patterns and trends of taxi use in a city are then identified and visually studied as taxi topics (clusters), thus relating street names and group behavior.

### B. *Mass mobility*

The works described in this subsection deal with analyzing people's collective mobility behavior, i.e., mass movements. This includes routine daily and weekly patterns as well as anomalies due to extraordinary events.

Von Landesberger et al. [49] present an approach to explore daily and weekly temporal patterns of collective mobility, where the source data are episodic trajectories of people reconstructed from georeferenced tweets or mobile phone use records. The trajectories are aggregated into flows between territory compartments by hourly intervals within the weekly time cycle. To reduce the complexity of the resulting set of flows, strongly connected neighboring compartments are aggregated into larger regions by means of density-based clustering. Then, similarly to Fig. 6, partition-based clustering of the time intervals according to the similarity of the spatial situations is used for revealing the periodic patterns of mass mobility. The situations corresponding to the time clusters are represented as graphs, i.e., node-link diagrams. Comparisons between clusters are supported by explicit visual encoding of the differences.

Beecham and Wood [19] present a technique for automatically identifying commuting behavior based on a spatial analysis of cyclists' journeys. They use visual analytics to compare the output of various workplace identification methods to explore data transformations and present insights to analysts in order to develop origin-destination theories of commute patterns. Ma et al. [54] also develop methods for studying urban flow. This work uses cell phone location records to approximate trajectories across a city, and flow volumes, links, and communities of users are visualized to help analysts identify typical patterns of movement within the city. Similarly, work by Yang et al. [78] focuses on identifying human mobility hotspots based on mobile phone location data from Shenzhen, China. Yang et al. applies kernel density estimation and clusters identified hotspots based on the temporal signatures to identify spatial locations with high travel demand.

Work by Chae et al. [22] develops a visual analytics framework for exploring public behavior before, during, and after disaster events. This work utilizes geographically referenced Tweets to create movement trajectories during disasters to identify evacuation flows. Interactions allow users to drill down into the data to also look at the underlying discourse occurring around the movements. Infrastructure data, disaster data (such as hurricane tracks), and Twitter data are all provided as map overlays in order to enable decision support and analysis.

### C. *People's activities and interests*

In order to understand the current use of transportation systems and plan for expansion and development, it is helpful to understand the reasons that people travel, i.e., the activities and interests related to traveling. This subsection reviews visual analytics works on transport-relevant knowledge discovery centered on people's use of space and reasons for traveling from population surveys and data obtained from social media.

Zhao et al. [80] visualize survey data concerning people's activities in space and time. Circular temporal histograms show the dependency of the activities on temporal cycles. A visualization technique called the ringmap is a variant of a circular histogram where aggregate values are shown by coloring and shading of ring segments. This allows aggregated data for different activities to be shown using multiple concentric rings.

Conducting population surveys is a costly and error-prone endeavor. Currently, due to the popularity of social media, researchers seek to obtain information about people's interests, activities, and purposes for traveling using social media mining. Photo sharing services, such as Flickr, have large numbers of georeferenced photos posted by people during their travels. Some of the photo posts have descriptive titles or tags indicating what attracted the photographers' attention. Other social media sites, such as Twitter, may have geo-coordinates embedded in the data, and recent works [4] [37] demonstrate the possibility of using these data for extracting information about people's interests in terms of places and events they like to visit.

To obtain semantic information related to people's mobility, researchers also explore other social media. For example, Krüger et al. [44] use data from Foursquare to attach semantic information to trips made with electric scooters. Specifically, they refer the trip origins and destinations to the categories of the places of interest located nearby, which may be indicative of the trip purposes. Andrienko et al. [4] explored the potential of georeferenced Tweets as a source of semantic information concerning people's activities and movements. They classified tweets according to the topics of the messages, such as 'food', 'coffee', 'education', 'sports', 'transportation', etc., and found that the topics corresponding to some activities tend to occur at the typical times of these activities. Thus, 'coffee' occurs mostly in mornings, 'food' at the lunch and dinner times, and 'sports' in the evenings and on the weekend. The authors also characterized different places regarding the topics that occur in the tweets posted in these places.

More recent work by Andrienko et al. [12] presents a procedure for obtaining data similar to personal daily mobility diaries. Such a diary reports what places were visited by a person during a day, at what times, and for what purposes. Mobility diaries from a large sample of population are a valuable source of information for transportation planning and simulation of various development scenarios. The presented procedure aims at extracting similar information from long-term sequences of spatio-temporal positions of people, which may come from georeferenced tweets or from mobile phone use records. From these sequences, the proposed procedure extracts repeatedly visited personal and public places along with the times these places were visited within the daily and weekly cycles. An interactive interface involving techniques for multi-criteria evaluation and ranking supports assignment of probable meanings ('home', 'work', 'eating', 'shopping', etc.) to subsets of places based on visit times and information about the land use or point of interest categories at these places. The analysis is done in a privacy-respectful manner without accessing individual data.

## V. Modeling and Planning

This section reviews research in visual analytics concerned with traffic modeling and transportation planning. This includes the derivation of models from data, applications of traffic forecasting and simulation models, transportation scheduling, and the exploration of decision options. We note that planning is also often done using retrospective data, and for such retrospective analysis, a variety of GIS tools and systems are well-established. Here, we focus on the use of forecasting models for large data as there are new, emerging challenges for visualization, interaction, and simulation at the intersection of intelligent transportation systems.

Scheepens et al. [61] describe two types of models that can be used for the prediction of individual movements of vessels. The first one is based on finding similar trajectories in a large historical database. The second model simulates the expected movement of a vessel based on its kinematic properties. Both models produce a prediction of the vessel positions over time as a temporal probability density field. The prediction is represented visually by contours showing the zones where the vessel is expected to be located at different times.

Sewall et al. [64][65] developed algorithms for simulation of movements of multiple vehicles in a street network. The

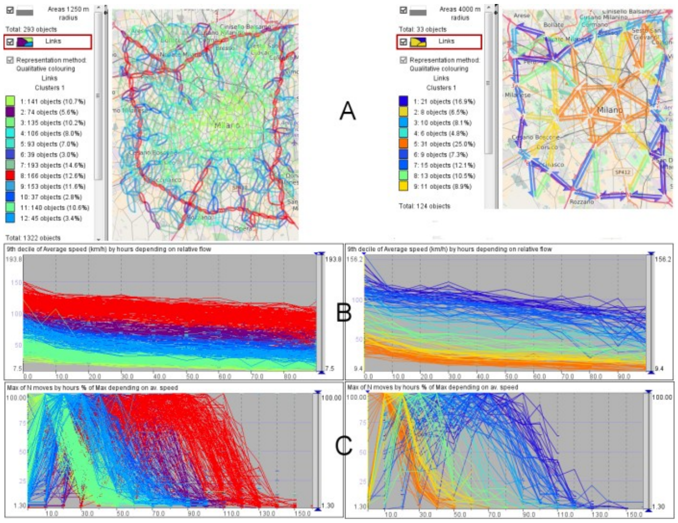


Fig. 9. Dependencies between the traffic flow intensities (hourly volumes) and mean velocities on the links of an abstracted transportation network at different levels of abstraction. A: Abstracted networks with the cell radii of about 1250 m (left) and 4000 m (right). The links are clustered and colored according to the similarity of the volume-speed dependencies. B: The dependencies of the mean velocity (vertical dimension) on the traffic flows (horizontal dimension): the velocities decrease as the flows increase. C: The dependencies of the flows (vertical dimensions) on the velocities (horizontal dimension): maximal flows can be achieved for certain velocities and decrease for both lower and higher velocities.[14]

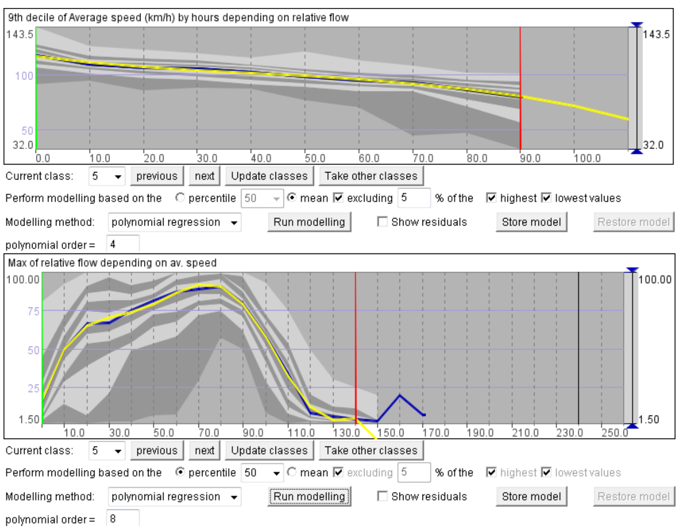


Fig. 10. For one of the clusters of links of an abstracted transportation network (see Fig. 8), the dependencies flow → velocity (top) and velocity → flow (bottom) are being represented by polynomial regression models.[10]

outputs are visualized as photorealistic 3D animations of the simulated traffic on selected junctions. One of the algorithms [65] involves a hybrid approach in which a detailed agent-based simulation of individual vehicle movements is done for user-selected areas of interest while a faster macroscopic model is used in the remainder of the network. There is an interactive interface for selecting regions to view in detail. In these works, the traffic forecasting is not based on previous analysis of historical data.

There is a series of works showing how predictive models of vehicle traffic can be derived from historical data consisting of a large number of vehicle trajectories [10][14][15]. The approach is based on spatial abstraction and aggregation of the

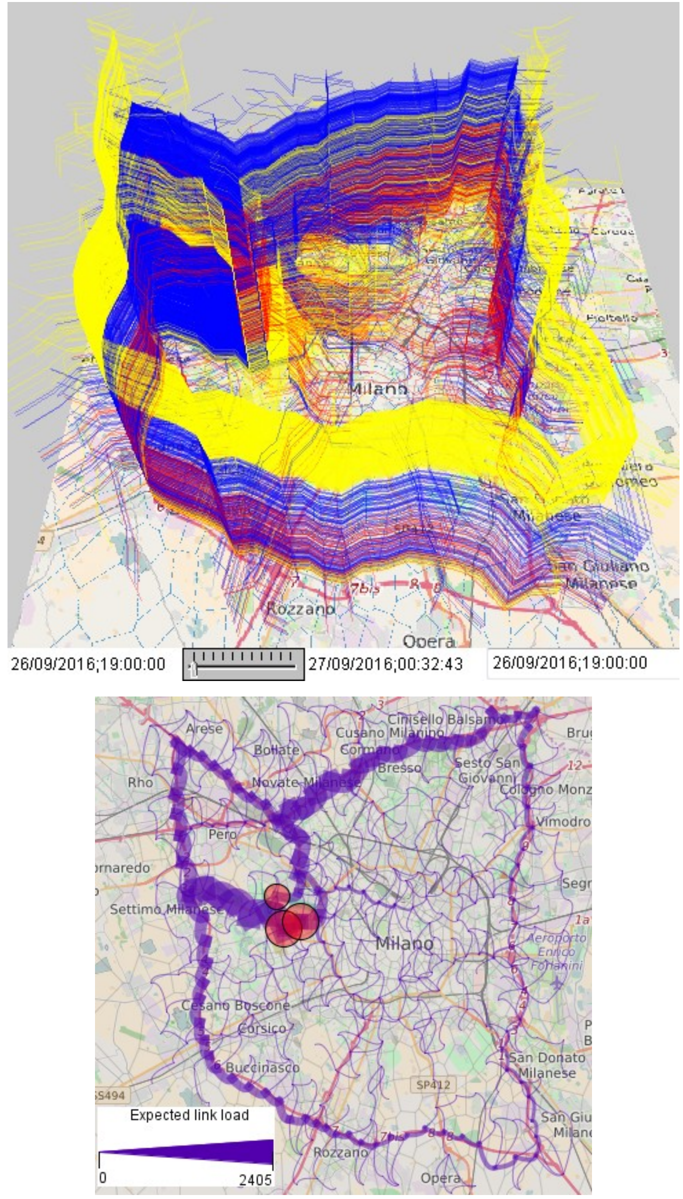


Fig. 11. Traffic flow – velocity dependency models extracted from historical traffic data (Figs. 8, 9) have been used for simulation of a scenario with 5,000 cars leaving the neighborhood of a stadium after a sport event. Top: The simulated trajectories of the individual cars are shown in a space-time cube. Bottom: the expected loads on the links of the abstracted traffic network are represented on the map by proportional widths of curved flow lines.[15]

trajectory data into collective movements (flows) of the vehicles between territory compartments, as shown in Fig. 1 (G, H). The authors discovered that the dependencies between the traffic intensities and mean velocities in an abstracted transportation network at different levels of abstraction (Fig. 9) have the same shapes as in the fundamental diagram of the traffic flow described in traffic theory [28]. While the fundamental diagram refers to links of a physical street network, it turns out that similar relationships also exist in abstracted networks. These dependencies can be represented by formal models (Fig. 10), which can be exploited to obtain fast predictions and simulations in cases when fine details are not necessary (Fig. 11).

Historical traffic data can be used not only for predicting future movements under various conditions but also for spatial planning applications. For example, the system SmartAdP [50] finds suitable locations for billboard placement using taxi trajectories. SmartAdP allows the user to select subsets of trajectories and areas of interest depending on the target audience and applies special algorithms for selecting optimal locations based on the traffic volumes and velocities. The system provides interactive visual tools for viewing, assessing, and comparing proposed candidate solutions.

The task of transportation scheduling is addressed by Andrienko et al. [3]. The general problem is to create a schedule for transporting a given set of items from their current locations to suitable destination places within a given time budget using an available fleet of transportation means. The items to transport may be of different categories requiring different kinds of transportation means. An example application is planning of evacuation of different groups of people, such as general population, schoolchildren, and hospital patients, from a disaster-affected area. The proposed system consists of a scheduling algorithm and a set of visual displays and interactive tools for exploring scheduling outcomes. The displays allow the user to detect problems, such as delays, understand their reasons, and find appropriate corrective measures.

## VI. TASKS AND DIRECTIONS FOR FURTHER RESEARCH

So far, most of the transportation-oriented research in visual analytics has been mostly focusing on exploring the opportunities created by the availability of huge amounts of mobility-related data, such as trajectories of vehicles, electronic records of the use of public transport, and digital traces of people using mobile devices. The visual analytics community has developed a solid knowledge base of the properties of these kinds of transportation-oriented data based on years of experience. Given that a large number of visual analytics methods, tools, and procedures have been developed, the exploratory mission of visual analytics for transportation-oriented data can thus be judged as quite successful.

As we noted in the introduction, visual analytics also has another mission: to find solutions to new problems or better solutions to old problems by using new opportunities. This mission of visual analytics is far from being fulfilled in the transportation domain. Only a few visual analytics works addressed specific transportation problems. Much of the transportation-related visual analytics research has been developed separately from transportation domain specialists and, hence, without proper knowledge of the domain problems. There is a clear need in closer cooperation between the communities and conducting interdisciplinary researches.

Future research should be mindful of emerging trends and continue seeking new opportunities. One possible direction is social transportation, which is a new concept that incorporates information from social space and cyber space into data acquired in physical space. Social media and mobile devices have recently experienced a rapid growth with the fast development of sensing, computing, and networking techniques. These social signals, from drivers' GPS coordinates, mobile phones' billing records to messages post on social media, record spatial, temporal and emotional information and establish the data foundation for social transportation research.

Furthermore, integrating analytics, interaction and novel visualizations with navigation systems can also be explored. As data on places of interest and personal preferences become available, navigation systems can incorporate such aspects to better inform drivers. For example, rerouting based on upcoming traffic situations can utilize metrics such as a driver's familiarity with a region. Data analyses should also work on incorporating more information about a city. While many works have looked at point of interest data and land parcel use, models and simulations of urban microclimate could also be incorporated to develop personal comfort routes during walking.

Along with developing visual analytics methods to support individuals in their travels, visual analytics should also focus on enabling modelers to interact with their simulations. As simulations get larger and more complex, new tools for exploring such high-dimensional data spaces are needed. This can be useful for simulating the future of mixed modality traffic (manned and driverless cars sharing the road), or for exploring how existing infrastructure can be efficiently updated. Such work will require combining heterogeneous data in a meaningful way, processing data 'in-vivo' and developing uncertainty-aware visualization techniques for spatiotemporal data, all of which are open challenges.

Finally, many of the tasks in transportation analysis require policy level decision making. As such, developing visual analytics tools that enable collaborative stakeholder engagement is critical. In this way, users can explore models and simulations, discuss the underlying assumptions, and inject real-world policy decisions into the models to explore potential future scenarios. Such collaborative visual analysis requires new tools, interactions, and data management systems.

## VII. CONCLUSION

Both the visual analytics community and transportation community has produced a large body of exploratory research work in analyzing transportation-related data. However, the knowledge acquired and methods developed often lack collaboration between the two communities. This overview and the special issue as a whole aim at raising the awareness of both visual analytics and transportation researchers and practitioners about the recent work in visual analytics, the essence of visual analytics approaches, and the high potential for solving complex problems that emerges from combining the power of computers with the unique capabilities of humans supported by interactive visual interfaces. We wish to build bridges between the visual analytics and transportation communities and promote their joint work for addressing various transportation problems.

## VIII. ACKNOWLEDGEMENTS

We greatly appreciate the advises on the content and organization of this paper received from M. Batty (University College, London, UK), J.J. van Wijk (Technical University of Eindhoven, the Netherlands), H. Schumann (University of Rostock, Germany), and P. Jankowski (San Diego State University, USA). We are also thankful to anonymous reviewers for constructive critiques and helpful suggestions.

## REFERENCES

[1] S. Al-Dohuki, F. Kamw, Y. Zhao, J. Yang, C. Ma, Y. Wu, X. Ye, F. Wang, X. Li, W. Chen. SemanticTraj: A New Approach to Interacting with Massive Taxi Trajectories. *IEEE Transactions on Visualization and Computer Graphics*. 23(1): 11-20, 2017.

[2] G. Andrienko, N. Andrienko, P. Bak, D. Keim, and S. Wrobel. *Visual Analytics of Movement*. Springer, 2013.

[3] G. Andrienko, N. Andrienko, U. Bartling. Visual Analytics Approach to User-Controlled Evacuation Scheduling. *Information Visualization*, 7(1): 89-103, 2008.

[4] G. Andrienko, N. Andrienko, H. Bosch, T. Ertl, G. Fuchs, P. Jankowski, D. Thom. Discovering Thematic Patterns in Geo-Referenced Tweets through Space-Time Visual Analytics. *Computing in Science and Engineering*, 15(3): 72-82, 2013.

[5] G. Andrienko, N. Andrienko, S. Bremm, T. Schreck, T. von Landesberger, P. Bak, D. Keim. Space-in-Time and Time-in-Space Self-Organizing Maps for Exploring Spatiotemporal Patterns. Computer Graphics Forum, 29(3): 913-922, 2010.

[6] G. Andrienko, N. Andrienko, G.Fuchs. Understanding Movement Data Quality. *Journal of Location Based Services*, 10(1): 31-46, 2016.

[7] G. Andrienko, N. Andrienko, G. Fuchs, J. Wood. Revealing Patterns and Trends of Mass Mobility through Spatial and Temporal Abstraction of Origin-Destination Movement Data. *IEEE Transactions on Visualization and Computer Graphics*. Accepted, 2016.

[8] G. Andrienko, N. Andrienko, C. Hurter, S. Rinzivillo, S. Wrobel. Scalable Analysis of Movement Data for Extracting and Exploring Significant Places. *IEEE Transactions on Visualization and Computer Graphics*, 19(7): 1078-1094, 2013.

[9] G. Andrienko, N. Andrienko, S. Rinzivillo, M. Nanni, D. Pedreschi, and F. Giannotti. Interactive Visual Clustering of Large Collections of Trajectories. In: *Proc. IEEE Conf. Visual Analytics Science and Technology*, pp. 3-10, 2009.

[10] N. Andrienko, G. Andrienko, A Visual Analytics Framework for Spatio-temporal Analysis and Modeling, *Data Mining and Knowledge Discovery*, 27(1): 55-83, 2013.

[11] N. Andrienko, G. Andrienko, Visual Analytics of Movement: an Overview of Methods, Tools, and Procedures. *Information Visualization*, 12(1): 3-24, 2013.

[12] N. Andrienko, G. Andrienko, G. Fuchs, P. Jankowski. Scalable and Privacy-respectful Interactive Discovery of Place Semantics from Human Mobility Traces. *Information Visualization*, 15(2): 117-153, 2016.

[13] N. Andrienko, G. Andrienko, G. Fuchs, S. Rinzivillo, H.-D. Betz. Detection, Tracking, and Visualization of Spatial Event Clusters for Real Time Monitoring. In: *Proc. IEEE/ACM Data Science and Advanced Analytics (DSAA)*, pp.1-10, 2015.

[14] N. Andrienko, G. Andrienko, S. Rinzivillo. Exploiting Spatial Abstraction in Predictive Analytics of Vehicle Traffic. *ISPRS International Journal of Geo-Information*, 4(2): 591-606, 2015.

[15] N. Andrienko, G. Andrienko, S. Rinzivillo. Leveraging Spatial Abstraction in Traffic Analysis and Forecasting with Visual Analytics. *Information Systems*, 57(1): 172-194, 2016.

[16] M. Batty. Editorial: Big Data, Cities and Herodotus. *Built Environment*, 42(3): 317-320, 2016.

[17] M. Batty. Big Data and the City. *Built Environment*, 42(3): 321-337, 2016.

[18] R Beecham, J Wood. Exploring gendered cycling behaviours within a large-scale behavioural dataset. *Transportation Planning and Technology*, 37 (1): 83-97, 2014.

[19] R. Beecham, J. Wood, A. Bowerman. Studying commuting behaviours using collaborative visual analytics. *Computers, Environment and Urban Systems*, 47:5-15, 2014.

[20] I. Boyandin, E. Bertini, P. Bak, D. Lalanne. Flowstrates: An approach for visual exploration of temporal origin-destination data. *Computer Graphics Forum*, 30(3): 971-980, 2011.

[21] J. Buchmüller, H. Janetzko, G. Andrienko, N. Andrienko, G. Fuchs, D.A. Keim. Visual Analytics for Exploring Local Impact of Air Traffic. *Computer Graphics Forum*, 34(3): 181-190, 2015.

[22] J. Chae, D. Thom, Y. Jang, S.Y. Kim, T. Ertl, D.S. Ebert. Public behavior response analysis in disaster events utilizing visual analytics of microblog data. Computers & Graphics, 38: 51-60, 2014.

[23] W. Chen, F. Guo, F.-Y. Wang. A survey on traffic data visualization. *IEEE Transactions on Intelligent Transportation Systems*, 16(6): 2970-2984, Dec. 2015.

[24] D. Chu, D.A. Sheets, Y. Zhao, Y. Wu, J. Yang, M. Zheng, G. Chen. Visualizing Hidden Themes of Taxi Movement with Semantic Transformation. In: *Proc. Pacific Visualization Symposium*, pp. 137-144, 2014.

[25] O. Ersoy, C. Hurter, F. Paulovich, G. Cantareiro, A. Telea. Skeleton-based edge bundling for graph visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12): 2364-2373, 2011.

[26] N. Ferreira, J. Poco, H.T. Vo, J. Freire, C. T. Silva. Visual exploration of big spatio-temporal urban data: A study of New York City taxi trips. *IEEE Transactions on Visualization and Computer Graphics*, 19(12): 2149-2158, 2013.

[27] A. Fredrikson, C. North C. Plaisant, B. Shneiderman, Temporal, geographic and categorical aggregations viewed through coordinated displays: a case study with highway incident data. In *Proc. Workshop on New Paradigms in Information Visualization and Manipulation (NPIVM'99)*, New York: ACM Press, 26-34, 1999.

[28] D.C. Gazis, *Traffic Theory*, Kliwer Academic, Boston, USA, 2002.

[29] D. Guo, X. Zhu, H. Jin, P. Gao, C. Andris. Discovering Spatial Patterns in Origin-Destination Mobility Data. *Transactions in GIS*, 16(3):411-429, 2012.

[30] H. Guo, Z. Wang, B. Yu, H. Zhao, and X. Yuan, TripVista: Triple Perspective Visual Trajectory Analytics and its application on microscopic traffic data at a road intersection. In: *Proc. Pacific Visualization Symposium*, pp. 163-170, 2011.

[31] K. Hamad and C. Quiroga, Geovisualization of Archived ITS Data-Case Studies, *IEEE Transactions on Intelligent Transportation Systems*, 17(1): 104-112, Jan. 2016.

[32] X. Huang, Y. Zhao, J. Yang, C. Ma, C. Zhang, X. Ye. TrajGraph: A Graph-Based Visual Analytics Approach to Studying Urban Network Centrality Using Taxi Trajectory Data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1): 160-169, 2015.

[33] E. van der Hurk, L. Kroon, G. Maróti and P. Vervest, Deduction of Passengers' Route Choices From Smart Card Data, *IEEE Transactions on Intelligent Transportation Systems*, 16(1): 430-440, Feb. 2015.

[34] C. Hurter, R.Alligier, D.Gianazza, S.Puechmorel, G.Andrienko, N.Andrienko. Wind parameters extraction from aircraft trajectories. *Computers, Environment and Urban Systems*, 47(1): 28-43, 2014.

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

[36] M. Itoh, D. Yokoyama, M. Toyoda, Y. Tomita, S. Kawamura, M. Kitsuregawa. Visual Exploration of Changes in Passenger Flows and Tweets on Mega-City Metro Network. *IEEE Trans. Big Data*, 2(1): 85-99, 2016.

[37] P. Jankowski, N. Andrienko, G. Andrienko, S. Kisilevich. Discovering Landmark Preferences and Movement Patterns from Photo Postings. *Transaction in GIS*, 4(6): 833-852, 2010.

[38] X. Jiang, C. Zheng, Y. Tian, R. Liang. Large-scale taxi O/D visual analytics for understanding metropolitan human movement patterns. *Journal of Visualization*, 18:185-200, 2015.

[39] C.R. Johnson, Top Scientific Visualization Research Problems. *IEEE Computer Graphics and Applications: Visualization Viewpoints*. 24 (4): 13–17, July 2004.

[40] D. A. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann, editors. *Mastering the Information Age – Solving Problems with Visual Analytics*. Eurographics Association, Goslar, 2010.

[41] L. M. Kieu, A. Bhaskar and E. Chung, Passenger Segmentation Using Smart Card Data, IEEE Transactions on Intelligent Transportation Systems, 16(3): 1537-1548, June 2015.

[42] C. Kinkeldey, A.M. MacEachren, M. Riveiro, and J. Schiewe,. Evaluating the effect of visually represented geodata uncertainty on decision-making: systematic review, lessons learned, and

recommendations. *Cartography and Geographic Information Science*, 44(1):1-21, 2017.
[43] M-J, Kraak and F. Ormeling *Cartography: visualization of spatial data*. Third edition. Guilford Publications, 2010.
[44] R. Krüger, D. Thom, T. Ertl. Semantic Enrichment of Movement Behavior with Foursquare – A Visual Analytics Approach. *IEEE Transactions on Visualization and Computer Graphics*, 21(8): 903-915, 2015.
[45] R. Krüger, D. Thom, M. Wörner, H. Bosch, T. Ertl. TrajectoryLenses – A Set-based Filtering and Exploration Technique for Long-term Trajectory Data. *Computer Graphics Forum*, 32: 451–460, 2013.
[46] P. A. Laharotte, R. Billot, E. Come, L. Oukhellou, A. Nantes and N. E. El Faouzi, Spatiotemporal Analysis of Bluetooth Data: Application to a Large Urban Network, IEEE Transactions on Intelligent Transportation Systems, 16(3): 1439-1448, June 2015.
[47] O.D. Lampe, H. Hauser. Interactive visualization of streaming data with kernel density estimation. In: *Proc. IEEE Pacific Visualization Symposium*, pp. 171-178, 2011.
[48] O.D. Lampe, J. Kehrer, H. Hauser. Visual Analysis of Multivariate Movement Data using Interactive Difference Views. In: Proc. Vision, Modeling, and Visualization Workshop VMV'2010, Siegen, Germany, pp. 315-322, 2010.
[49] T. von Landesberger, F. Brodkorb, P. Roskosch, N. Andrienko, G. Andrienko, A. Kerren. MobilityGraphs: Visual Analysis of Mass Mobility Dynamics via Spatio-Temporal Graphs and Clustering. *IEEE Transactions on Visualization and Computer Graphics*, 22(1): 11-20, 2016.
[50] D. Liu, D. Weng, Y. Li, J. Bao, Y. Zheng, H. Qu, Y. Wu. SmartAdP: Visual Analytics of Large-scale Taxi Trajectories for Selecting Billboard Locations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1): 1-10, 2017.
[51] H. Liu, Y. Gao, L. Lu, S. Liu, H. Qu, L.M. Ni. Visual analysis of route diversity. *In Proc. 2011 IEEE Conf. Visual Analytics Science and Technology (VAST)*, pp. 169-178, 2011.
[52] S. Liu, J. Pu, Q. Luo, H. Qu, L.M. Ni, R. Krishnan. VAIT: A Visual Analytics System for Metropolitan Transportation. IEEE Transactions on Intelligent Transportation Systems, 14(4): 1586-1596, 2013.
[53] P. Lundblad, O. Eurenius, T. Heldring, Interactive Visualization of Weather and Ship Data. In*: Proc. 13th Int. Conf. Information Visualization IV2009*, IEEE Computer Society Press, pp. 379-386, 2009.
[54] Y. Ma, T. Lin, Z. Cao, C. Li, W. Chen. Mobility Viewer: A Eulerian Approach for Studying Urban Crowd Flow. *IEEE Transactions on Intelligent Transportation Systems*. To appear, 2016.
[55] C. Palomo, Z. Guo, C.T. Silva. Visually Exploring Transportation Schedules. *IEEE Transactions on Visualization and Computer Graphics*, 22(1): 170-179, 2016.
[56] J. Pu, S. Liu, Y. Ding, H. Qu, L. Ni. T-Watcher: A new visual analytic system for effective traffic surveillance. In *Proc. IEEE 14th Int. Conf. Mobile Data Management*, pp. 127-136, 2013.
[57] Y. Qiang, S.H. Chavoshi, S. Logghe, P. De Maeyer, N. Van de Weghe. Multi-scale Analysis of Linear Data in a Two-Dimensional Space, *Information Visualization*, 13(3): 248-265, 2014.
[58] S. Rinzivillo, D. Pedreschi, M. Nanni, F. Giannotti, N. Andrienko, G. Andrienko. Visually–driven analysis of movement data by progressive clustering. Information Visualization, 7(3/4): 225-239, 2008.
[59] R. Scheepens, C. Hurter, H. van de Wetering, J.J. van Wijk. Visualization, selection, and analysis of traffic flows. IEEE Transactions on Visualization and Computer Graphics, 22(1): 379-388, 2016.
[60] R. Scheepens, S. Michels, H. van de Wetering, J.J. van Wijk. Rationale Visualization for Safety and Security, *Computer Graphics Forum*, 34(3): 191-200, 2015.
[61] R. Scheepens, H. van de Wetering, J.J. van Wijk. Contour based visualization of vessel movement predictions. *International Journal of Geographical Information Science*, 28(5): 891-909, 2014.
[62] R. Scheepens, H. van de Wetering, J.J. van Wijk. Non-overlapping aggregated multivariate glyphs for moving objects. In: Proc. Pacific Visualization Symposium (PacificVis), IEEE, pp.17-24, 2014.
[63] R. Scheepens, N. Willems, H. van de Wetering, G. Andrienko, N. Andrienko, J. J. van Wijk. Composite Density Maps for Multivariate Trajectories. IEEE Transactions on Visualization and Computer Graphics, 17(12): 2518-2527, 2011.
[64] J. Sewall, D. Wilkie, P. Merrell, M.C. Lin, Continuum Traffic Simulation, *Computer Graphics Forum*, 29(2): 439-448, 2010.
[65] J. Sewall, D. Wilkie, and M.C. Lin, Interactive Hybrid Simulation of Large-Scale Traffic, *ACM Transactions on Graphics*, 30(6), Article 135, 2011.
[66] G. Sun, R. Liang, H. Qu, Y. Wu. Embedding Spatio-temporal Information into Maps by Route-Zooming. *IEEE Transactions on Visualization and Computer Graphics*. Accepted, 2016.
[67] J.J. Thomas and K.A. Cook, editors. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, IEEE CS Press, 2005.
[68] W. Tobler, Experiments in Migration Mapping by Computer, *The American Cartographer*, 14(2): 155-163, 1987.
[69] C. Tominski, H. Schumann, G. Andrienko, N. Andrienko. Stacking-Based Visualization of Trajectory Attribute Data. *IEEE Transactions on Visualization and Computer Graphics*, 18(12): 2565-2574, 2012.
[70] F. Wang, W. Chen, F. Wu, Y. Zhao, H. Hong, T. Gu, L. Wang, R. Liang, H. Bao. A Visual Reasoning Approach for Data-driven Transport Assessment on Urban Road. In: *Proc. IEEE Conf. Visual Analytics Science and Technology VAST 2014*, pp. 103-112, 2014.
[71] Z. Wang, M. Lu, X. Yuan, J. Zhang, H. van de Wetering. Visual traffic jam analysis based on trajectory data. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2159–2168, 2013.
[72] Z. Wang, T. Ye, M. Lu, X. Yuan, H. Qu, J. Yuan, Q. Wu. Visual Exploration of Sparse Traffic Trajectory Data. *IEEE Transactions on Visualization and Computer Graphics*, 20(12): 1813 – 1822, 2014.
[73] N. Willems N, H. van de Wetering, J.J. van Wijk, Visualization of vessel movements. *Computer Graphics Forum*, 28(3): 959-966, 2009.
[74] K. Wongsuphasawat, M.L. Pack, D. Filippova, M. VanDaniker, A. Olea. Visual Analytics for Transportation Incident Data Sets. *Transportation Research Record: Journal of the Transportation Research Board*, 2138: 135-145, 2009.
[75] J. Wood, J. Dykes, A. Slingsby. Visualisation of origins, destinations and flows with OD maps. *The Cartographic Journal*, 47(2): 117 – 129, 2010.
[76] J. Wood, A. Slingsby, J. Dykes. Visualizing the dynamics of London's bicycle hire scheme. *Cartographica*, 46(4): 239 – 251, 2011.
[77] M. Wörner, T. Ertl, Visual Analysis of Public Transport Vehicle Movement. In: *Proc. Int. Workshop on Visual Analytics (EuroVA 2012)*, pp. 79-83, 2012.
[78] X. Yang, Z. Zhao, S. Lu. Exploring spatial-temporal patterns of urban human mobility spots. *Sustainability*, 8(7): 674, 2016.
[79] W. Zeng, C. W. Fu, S. M. Arisona, A. Erath, H. Qu. Visualizing Mobility of Public Transportation System. *IEEE Transactions on Visualization and Computer Graphics*, 20(12): 1833-1842, 2014.
[80] J. Zhao, P. Forer, A.S. Harvey. Activities, ringmaps and geovisualization of large human movement fields. *Information Visualization*, 7(3): 198-209, 2008.
[81] X. Zhu, D. Guo. Mapping Large Spatial Flow Data with Hierarchical Clustering. *Transactions in GIS*, 18(3):421-435, 2014.

Gennady Andrienko is a lead scientist responsible for the visual analytics research at the Fraunhofer Institute for Intelligent Analysis and Information Systems (IAIS) and professor (part-time) at City University London. He co-authored two monographs 'Exploratory Analysis of Spatial and Temporal Data' (Springer, 2006) and "Visual Analytics of Movement" (2013) and 80+ peer-reviewed journal papers. From 2007 till 2015 Gennady Andrienko was chairing the ICA Commission on GeoVisualization. He co-organized scientific events on visual analytics, geovisualization and visual data mining, and co-edited 13 special issues of journals.

Natalia Andrienko has been working at GMD, now Fraunhofer IAIS, since 1997. Since 2007, she is a lead scientist responsible for the visual analytics research. Since 2013 she is professor (part-time) at City University Londond. She co-authored the monographs “Exploratory Analysis of Spatial and Temporal Data” (Springer, 2006) and ”Visual Analytics of Movement” (Springer, 2013) and over 70 peer-reviewed journal papers. She received best paper awards at AGILE 2006, EuroVis 2015 and IEEE VAST 2011 and 2012 conferences, best poster awards at AGILE 2007, ACM GIS 2011 and IEEE VAST 2016, and VAST challenge awards 2008 and 2014.

Dr. Wei Chen is a professor in State Key Lab of CAD&CG at Zhejiang University, P.R.China. He has performed research in visualization and visual analysis and published more than 30 IEEE/ACM Transactions and IEEE VIS papers. His current research interests include visualization, visual analytics and bio-medical image computing. He served as paper co-chairs of IEEE Pacific Visualization 2013 and ACM SIGGRAPH Asia 2016 Symposium on Visualization.

Ross Maciejewski is an Assistant Professor at Arizona State University. His recent work has actively explored the extraction and linking of disparate data sources exploring combinations of structured geographic data to unstructured social media data to enhance situational awareness. His primary research interests are in the areas of geographical visualization and visual analytics focusing on public health, dietary analysis, social media, and criminal incident reports. He is a Fellow of the Global Security Initiative at ASU and the recipient of an NSF CAREER Award (2014).

Ye Zhao is an associate professor in the Department of Computer Science at the Kent State University, Ohio, USA. His current research projects include visual analytics of urban transportation data, multidimensional, text, and animated information visualization, and computational hemodynamics modeling. He has co-authored more than 50 refereed technical papers and served in many program committees including IEEE SciVis and VAST conferences. His work has been supported by research grants from NSF, Ohio Board of Agents, etc. He received Google Faculty Research Award in 2011.