Position and Movement Analytics

Situation

With the emergence of PNDs (personal navigation devices), use of GPS-enabled smart phones, and deployment of other location sensitive devices, a tsunami of real-time data describing position and movement of assets, resources and people is becoming available. This new class of data consists of streams of time-stamped locations generated with attributes at each location. The attributes might be the speed, acceleration, or engine temperature of a vehicle. More complex and interesting to business efficiency and effectiveness attributes include the temperature of a cold storage container sampled through time while in route through the supply chain, opinion of a panelist responding to a real-time survey on a mobile phone, the number of times a mobile subscriber has visited a particular store, or even the number of riders on a bus sampled at every stop. What is common about this class of data is that sequence of time-stamped records, which we call *Position Facts*, are delivered over wireless networks as a stream.

Opportunity for Real-time Analytics

There is an opportunity to use analytics to process this data stream to create business value. Looking at this data in aggregate, statistics can be computed that can generate probabilities, or correlations that support decision making or action taking. For example, in a mobile marketing application the particular offer sent to a subscriber (who opts in) might depend on the previous sequence of stores that she has passed, her age, and her previous purchasing patterns. A trucking company might want to analyze the variability in delivery times, correlate variability with driver and load, and then use the results to identify friction points in movement patterns. A social media company might be interested in relating movement and buying patterns to determine the optimum offer for both advertisers and individuals. A logistics company needs to analyze movement patterns to optimize the flow of goods and ensure on-time deliveries. A telecommunications company might wish to score their customer base to see which subscribers are at risk of churning.

The massive quantity of real-time Position Facts is overwhelming, thereby making it challenging, if not impossible, to extract deep value form the data using traditional tools and approaches. The business need is to create smart SaaS solutions that produce this deeper value. Using the VisTracks platform, Position and Movement Analytics applications process streams of data in real-time and enable the processed position and movement information to be analyzed and correlated with other data. What is common about these applications is that they involve computing on Position Facts, performing statistical analyses, making predictions, enabling decisions and taking real-time actions.

Solution – VisTracks Analytics Platform

To respond to the opportunity, VisTracks has developed a multi-tenant, cloud-based *Position and Movement Analytics*™ platform for solution providers who need to develop software for sensing, recording, reacting, analyzing, and predicting – from real-time, and ongoing location-specific data streams. The VisTracks system is a platform-as-a-service (PaaS) that provides programmable web services for advanced statistics, tracking, rules-based fact processing and notification and visualizing spatially relevant analysis of data. When compared with other solutions, VisTracks analytics are richer and enable SaaS developers to use the complete history of movements in their solutions. For example, many location-based services depend just on the current location of the mobile subscriber and not profiles of her movement patterns.

There are seven components to our analytics platform.

1. Real-time Data Warehouse – Data Asset

As new position facts are processed by the platform they are organized into a multi-tenant position and movement cloud-based data warehouse specifically engineered for manipulating this new class of data. Our data warehouse currently contains hundreds of millions of position facts collected from moving entities that we analyze to identify movement patterns. The position facts and aggregate data structures are stored in a cloud-based data warehouse.

Through time as more and more devices are provisioned onto our platform we anticipate creating a valuable data asset. We proper ammonization VisTracks will be able to extract value from this data by mining it to understand movement patterns, segmenting the objects in the data warehouse into clusters, etc.

2. Position and Movement Data Structures – New Schema

The focus of the VisTracks platform is to provide SaaS developers with APIs, data structures, computational algorithms, and predictive statistics to facilitate the easy creation of rich solutions. As new Position Facts arrive, the VisTracks platform maintains and updates internal data structures that are specifically tailored to this class of data. The platform provides APIs that provides solution providers with easy access to monitor, query and act on the information. At a high level, the VisTracks data structures include:

1. **Position Facts and Attributes.** After a tracking device, smart phone, or other type of sensor is provisioned on the platform, the system begins listening for and recording all incoming reports from the device. The resulting history, representing the complete state of the device through time, is available to application developers through a set of APIs, along with all associated attribute values. These APIs make it easy for developers to create real-time tracking solutions, and identify if the sensor or mobile phone is in a location where it should receive a pre-programmed message.

The use case for providing the history of movement and attribute values to the developer through APIs instead of just the current position is that it enables the solutions to solve more complicated business problems. For example, the pre-programmed message can depend on the sequence of positions and attributes leading to the current position.

2. **Trajectories.** A trajectory is a data structure that captures the historical movements of an asset. A trajectory begins at a *Stop Point*, a location where an object has been stationary for a fixed time interval and movement distance threshold, e.g. 15 minutes and 25 meters. The reason for the distance threshold is to account for jittering and randomness that occur in successive GPS readings. A trajectory also moves continuously through a sequence of positions, and then ends at another stop point. The platform automatically calculates trajectories and stop points as each new Position Fact arrives. The attributes from the Position Facts are associated with each position along the trajectory. Also, the platform calculates various statistics on the trajectories such as average, minimum, and maximum speeds, distance travelled, etc. The platform provides APIs for calculating various statistics along a trajectory for attributes. For example, the statistics might include the average humidity, condition, state, temperature, pressure, weight, and shrinkage through time.

   The use case for trajectories is to understand detailed continuous movement patterns. One of our APIs enables a SaaS developer to extract all trajectories that passed within a specified distance of well-known location. For a marketing application, a SaaS developer could know how often a consumer with a tracked smart phone passed within ½ mile of a store location and the time of day that it went past. These consumers might be unaware of the store's location and would be prospective and could be marketed to.

   The business ROI enabled by Trajectories is that they enable analysis of movement patterns, trending, and friction points. Simple tracking systems just monitor the location of assets and do not capture the history of movements.

3. **Trips.** Multiple trajectories combine to make the concept of trips. For example, the path taken by a commercial vehicle driving from Montreal to New York will consist of a sequence of trajectories. The first trajectory in the trip might be the path from the garage where the vehicle was stored overnight to the pickup location for the shipment. The second trajectory is the path from the loading dock to the first rest stop for the driver, etc.

   The use case for trips is trips are the natural way to think about supply chain problems. A commercial vehicle making a deli very from Canada to the USA would make a single trip consisting of several trajectories.

4. **Stop Points.** Stop points are locations where the moving entity was stationary for a fixed period. There is a settable threshold, e.g., 15 minutes, which is used by the platform to

determine whether a short stop is really temporary pause in the movement, such as might be caused by traffic, or a true stop point.

Stop points are useful to understand pauses in movement, which often represent friction points. A pattern of excessively long stop points at a delivery location may indicate a loading and unloading inefficiency that needs to be addressed.

5. **Position and Movement Signatures.** For each object, the VisTracks platform calculates a movement signature that characterizes the movement patterns of the object. The movement signatures are probabilistic representations of the movement patterns and can be used to classify the different types of movement. For example, a movement signature that shows a week day pattern of movement from the Chicago suburbs downtown and back is characteristic of a commuter.

   The concept of signatures is more valuable than traditional threshold-based analysis because a signature incorporates many dimensions of movement and there allows more flexibility in categorizing and predicting movement patterns. Conversely, a threshold-based classification is inherently one dimensional and therefore quite limited. For example, one use case for movement signatures is to identify patterns of excessively aggressive driving, which might indicate an insurance risk. Aggressive driving can only be identified in context of the situation when considering multiple dimensions including the driving patterns of the other vehicles, maximum speed, acceleration, fast lane changes, hard braking events, high speeds during turns, fast decelerations, etc. A simple threshold-based system cannot accurately identify aggressive driving because it does not include the full context of the activity.

6. **Pattern Libraries.** Pattern libraries are useful to organize and segment movement patterns. For example, one type of pattern library might describe movement patterns for daily commuters. Another type might describe the movement patterns of college students or perhaps young stay-at-home mothers.

The VisTracks platform extends traditional spatial databases by creating new types for position and movement data. In a spatial database, the base types are points, lines, polygons, shapes, boundaries, etc. The VisTracks platform extends a spatial database by adding new base types for trajectories, stop points, trips, signatures, etc. These new types are first class database objects, offering a complete set of methods, properties, and algorithms for manipulating the types, extracting information, making comparisons, etc.

For example, one VisTracks extension involves the trajectory type. By making a trajectory a first class object, it becomes possible to compare movements statistically. E.g. for trajectories representing movements between Lisle, Illinois and Chicago, we can correlate the mean and standard deviation of the trajectory duration with the time of day, and thereby set confidence intervals on expected delivery times.

3. Position and Movement Analytics (First Order Analytics)

For many applications there is a need for analytics around the data structures in the VisTracks data warehouse. These analytics might involve simple sorting and grouping operations and are implemented by adding primates to the spatial database. Examples of the operations include:

1. **Routing and Distance Analytics** – what is the driving distance (walking) between locations? Which mobile object is available and closest to an unexpected incident? What is the average speed, driving time, distance, etc. for objects in a fleet, mobile assets, objects being dispatched?

2. **Trip and Trajectory Analytics.** The use case is to determine which objects have passed by a designated location or set of locations during a specified time period. What is the average speed of a trip? What trips represent efficient movement patterns? Which are inefficient? Which trips have predictable movement patterns?

3. **Stop Point Analytics.** For mobile resources it is important to understand both movement and stop patterns. E.g. which stops take the most time? Is the vehicle stopped when it could be moving? Do stop points come at the edge of the day? Could schedules be rearranged to minimize stop points? What is the ROI of shortening a stop point? Is the vehicle left idling during a stop point?

4. **Correlate Movements with Other Attributes.** When other attributes such as idle time, vehicle speed, hard braking events, delivery delays, etc. are available there is a need to correlate this information with movement data.

4. Statistical and Data Mining Algorithms (Second Order Analytics)

The platform contains several types of statistical algorithms. As new position facts arrive, the analytics engine automatically analyzes the data, characterizes movement patterns, segments the movements, and adaptively learns from the data. For each position and movement data type the platform provides statistical algorithms to understand variability and predictability of movement. The platform provides statistical algorithms that quantify predictability and variability of movement by time of day, day of week, location, route, etc.

1. **Variability and Predictability.** The VisTracks platform provides statistical algorithms to understand variability and predictability of movement. For a Canadian trucking company making just-in-time deliveries to a Detroit auto manufacturer it is important to be able to accurately predict border crossing times for the Ambassador bridge to ensure accurate delivery times. The platform provides statistical algorithms that quantify predictability and variability of movement by time of day, day of week, location, route, etc.

2. **KPIs and Indices to measure efficiency.** There is a need to develop metrics, KPIs, and indices that measure movement efficiency. The first order analytics capture raw movements. The second order analytics are KPIs and indices that determine if the movement patterns are efficient. For

example, which transportation corridors are the most efficient for shippers? What KPIs measure the performance a fleet with respect to its competitors? With respect to the best efficiency that could be achieved?

3. **Clustering to Identify Movement Patterns with Collaborative Filtering**. By clustering objects with similar position and movement patterns we can segment the objects into similar groups. For example, one cluster of objects might correspond to young professionals who spend weekend evening hours in trendy night life locations and live in fast-moving diverse communities. Another cluster might spend evening hours in the Latino section of Chicago and in middle class neighborhoods. Using collaborative filtering, e.g. what positions and movements are common to others in an object's cluster, the VisTracks platform can suggest locations that might be of interest to an object.

4. **Segmentation**. Clustering naturally leads to segmentation. Advertisers want to segment their prospect database to provide personalized offers. The segments will use movement history and demographic variables to provide meaningful groupings for personalized marketing.

5. **Scoring**. Based on movement patterns, the platform provides the capability to fit a predictive analytics model to score objects for a capability. The algorithm works by fitting the predictive model to a training set and then applying the algorithm to the complete object set.

These algorithms are implemented by integrating a statistical engine and APIs for predictive analytics with the VisTracks data warehouse and schema. As a result, it is possible to perform regressions, correlation analysis, and make stochastic predictions using statistical models.

5. Programmable Rules Engine for Complex Event Processing

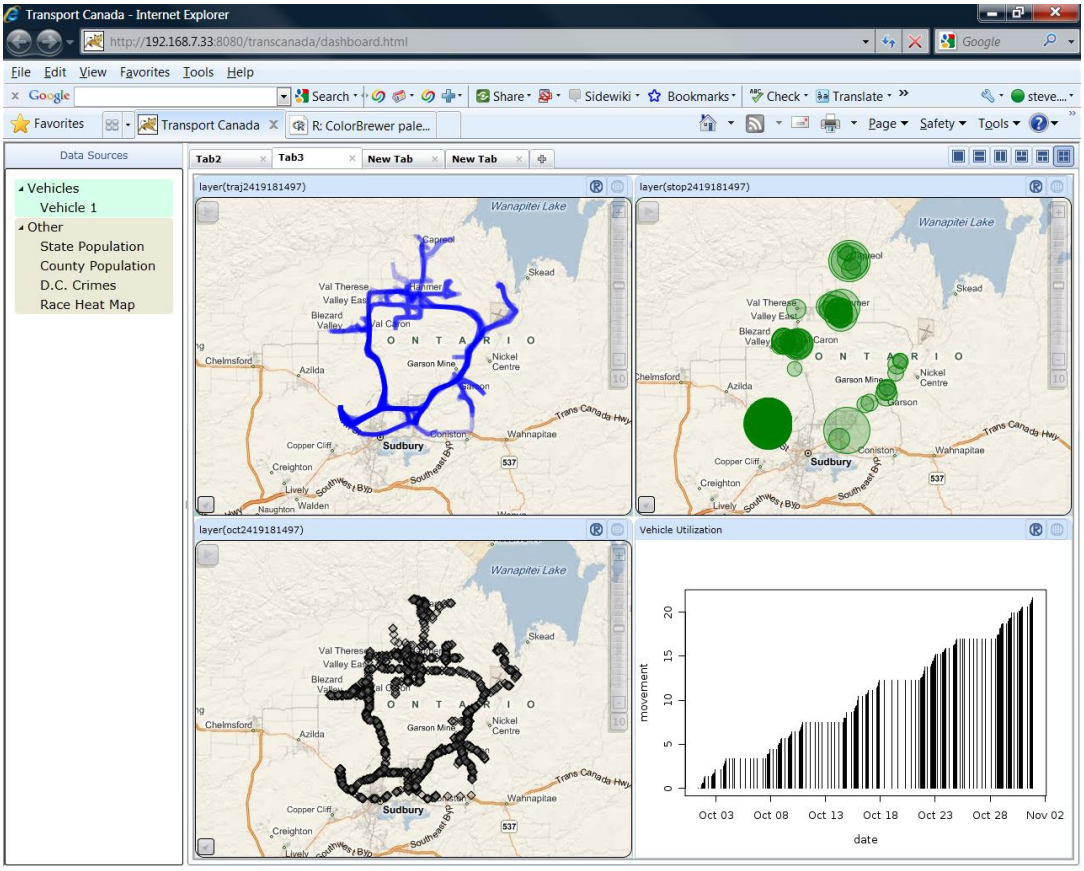
The platform provides a programmable rules engine with complex logic that enables developers to create guided or arbitrary rules to create actions. For example, the actions might be to stand up an alert, send a mobile subscriber a SMS, trigger an alarm, send an email message, generate a list, etc.

In contrast with the rules engine for a simple tracking solution or location-based services, what is unique about the VisTracks rules engine is that it supports complex event processing on the full history of movement of an entity. A complex rule, for example, might involve which POIs the trajectory of an entity passed, its speed when it passed the POIs, time of day, etc.

The platform, which uses the Drools Rules Engine (Drools Business Logic Integration Platform, 2010), provides two ways to implement a rule. First, a rule can be constructed using the raw Drools Rule Language (Wikipedia, Introduction to Drools Rule Language, 2010). Using this method an arbitrarily complex rule can be created. Second, VisTracks provides a way to create a *Rule Template*. A rule template is effectively a parameterized raw rule. The rule template is the instantiated as a *Guided Rule* to create a running instance of the rule. Guided rules combine a rule template with parameters to create an instance of a rule that executes on the platform.

6. Real-time Visual Analytics Dashboard

VisTracks provides a configurable visual analytics dashboard as a dynamic presentation interface to do quick visual analysis of position and movement data. The dashboard is fully configurable, real-time, and enables customers to quickly access statistics associated with the most recent information streaming into the system. The dashboard includes maps, timelines, bar charts, scatterplots, and other interactive tools to support ad hoc analysis.



VisTracks real-time analytics dashboard.

7. Platform APIs

The VisTracks platform APIs are the mechanism that developer s use to create applications using the VisTracks platform. There are several classes of APIs

1. **Simple APIs** – these are commodity APIs that all platforms must provide to support basic operations.

- Routing – generate the shortest, quickest, or other constraint-based route between two locations
- Geocoding – return the location of an address
- Reverse-geocoding – for a location, return the address
- Spatial Data Services – manage map content, simple polygons for geofences, etc.

2. **Tracking APIs** – provides capability to build real-time tracking applications.
- Current location of all objects
- Movement history (trajectory) – the sequence of positions of an object over any time period
- Geofences for alerts

3. **Position and Movement Analytics (First Order) –** this set of APIs provides analytical functions on VisTracks core VisTracks data types (Trajectories, Trips, Stop Points, …)
- Statistics on Trips/Trajectories/Stop Points – e.g. distance, time, average, shortest, longest, …
- Distribution of trip times between two city pairs
- Ratio of Stop Point times to total trip times
- Analysis of Arrival/Departure Times
- Average Speeds by time of day, day of week
- Numbers of acceleration and deceleration events if this attribute is provided by the device
- Numbers of hard braking events if this attribute is provided by the device
- Idle time, this attribute is provided by the device

4. **Signatures and Movement Pattern APIs –** this set of APIs is for extracting signatures and creating pattern libraries.
- Retrieve movement signature for an object
- Create/delete pattern
- Create pattern library
- Delete pattern library
- Duplicate pattern
- Show differences in movement patterns

5. **Statistical and Data Mining APIs (Second Order Analytics)**
- Measure predictability and variability of object movement
- Create KPIs which determine efficiency of movement in comparison with the most efficient object movement patterns
- Cluster set of objects based on movement patterns
- Retrieve top n locations for a set of objects
- Segment the objects into homogeneous categories with similar position and movement patterns
- Score objects for an offer, e.g. for a SMS coupon, for a market research survey, ….

6. **Rules Engine APIs –** this set of APIs enables a developer to manipulate the VisTracks rules engine
- Arbitrary Rules based on JBoss Rules reasoning engine
- Rule Templates that create parameterized rules
- Guided rules that are instantiated instances of the rules templates

7. **Visualization APIs –** these APIs enable the developer to generate

- Create heat map of object positions at specified time
- Create a business graphics charts (bar/pie/lines, …)
- Create distribution of trip times as a statistical display
- Visualize selected position and movement variables on a map

Conclusion

VisTracks has created a next generation cloud-based PaaS system for solution providers to create real-time position and movement applications. The platform extends spatial databases by providing new data structures for position and movement data and the APIs to develop powerful solutions. These data structures capture the complete movement patterns and attribute values of an object. In addition, the platform includes statistical computations that characterize movement patterns and support change detection.