

# The Tale of (Fusing) Two Uncertainties

Bing Zhang\*  
Northwestern University  
2145 Sheridan Road  
Evanston, IL, 60201  
bing@u.northwestern.edu

Goce Trajcevski†  
Northwestern University  
2145 Sheridan Road  
Evanston, IL, 60201  
goce@eecs.northwestern.edu

## ABSTRACT

This work addresses the problem of fusing spatio-temporal uncertainties obtained from heterogeneous location sources: on-board GPS devices and roadside sensors. We develop a model for combining the uncertain location-values from the different sources, which further narrows the possible locations of a given object. Our experiments demonstrate that the proposed model may eliminate significant amount of the false positives, compared to the traditional space-time prism (bead) uncertainty models.

## Categories and Subject Descriptors

H.2.8 [Database Applications]: Spatial databases and GIS

## General Terms

Theory

## Keywords

Uncertainty Fusion, Roadside Sensors, Beads

## 1. INTRODUCTION

The volume of location-in-time data exceeds the order of Peta-Bytes per year just from smartphones [5]. Many applications relying on Location Based Services (LBS) rely on efficient techniques for storage, retrieval and query processing for such data – topics studied in the field of Moving Objects Databases (MOD) [4]. Due to the inherent imprecision of the sensing devices, typically there is a degree of *uncertainty* associated with the measurements/values. The problem of capturing the uncertainty into the data-models [7] as well as queries' syntax and processing algorithms [12, 11] has been recognized and tackled by several earlier works.

\*Research Supported by the NSF III 1213038 grant

†Research Supported by the NSF grants CNS 0910952 and III 1213038, and ONR grant N00014-14-10215

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.  
SIGSPATIAL '14, November 04 - 07 2014, Dallas/Fort Worth, TX, USA  
Copyright 2014 ACM 978-1-4503-3131-9/14/11 ...\$15.00  
<http://dx.doi.org/10.1145/2666310.2666495>

Complementary to obtaining location by on-board Global Positioning System (GPS) device – typical for MOD settings – in many traffic management applications [1] the location data is obtained from some types of road-side sensors. For example, lane level positioning is an important component in navigation systems widely applied in smart traffic control, automated vehicle location or intelligent transportation systems [10]. Such sensors-data may be combined with data from different on-board sensing devices – e.g., U.S. Xpress gathers 900 to 970 data elements of various engine/component readings [8].

The motivation for this work is the observation that, to date, there has been no model that combines the uncertain data from two different sources. Given a query:

**Q1<sup>u</sup>:** *Retrieve all the vehicles which have had  $> \Theta$  ( $0 < \Theta \leq 1$ ) probability of crossing the lane in road segment RS1 when driving less than 50km/h and carrying less than 80% of the maximum load.*

we argue that properly considering the joint impact – equivalently *fusion* – of the uncertainties from the GPS sources and road-side sensors can eliminate some of the moving objects (trajectories) from its answer-set.

Our main contribution is a novel model of spatio-temporal uncertainty for moving objects, which combines the location data obtained by GPS devices on-board moving objects and the location data obtained from road-side sensors. We also report our preliminary experimental observations, demonstrating the reduction of false positives from the answers to certain spatio-temporal queries.

## 2. PRELIMINARIES

We now overview some techniques for location data – both GPS-based and based on road-side sensors. Subsequently, we proceed with introducing the basic notation used in the rest of the paper.

Starting in the 1920s, when the traffic signals were still manually controlled, several generations of sensor types have been developed and deployed for traffic management – from pressure-sensitive sensor in 1931 to modern laser sensors [2].

Contrary to the GPS-based data acquisition techniques where each data source is isolated, the roadside sensors are usually connected hierarchically to a server and send their sampled data to traffic control center [6]. Compared with GPS system, the roadside sensors have better measurement accuracy, higher sampling frequency and shorter response time, which enables their use in real time traffic information analysis and control.

When it comes to location obtained via on-board GPS

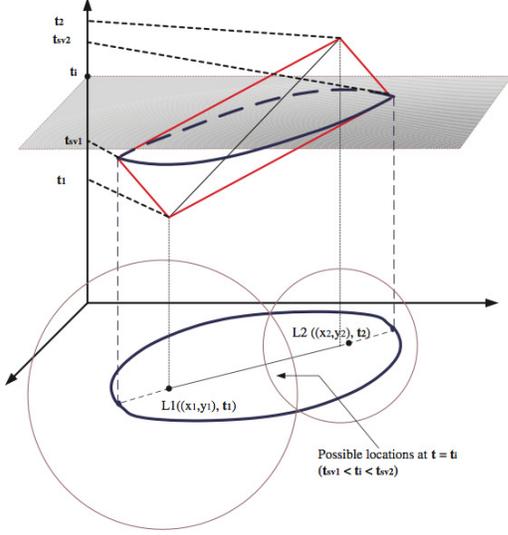


Figure 1: Bead and Ellipse Model

devised, a basic approach for modeling spatio-temporal uncertainty of moving objects is the, so called, sheared cylinder model. The main assumption is that at any time instant  $t_i$ , the object's location is inside a given disk with a fixed radius, centered at the expected location at  $t_i$ . For time values different from sampling ones, the expected location is obtained via linear interpolation [12]. However, this model is geared towards past/historic trajectories.

The implications of the fact that the object's motion was bound by some  $v_{max}$  in-between two updates was analyzed in [9]. Based on the definition as a geometric set of 2D points, it was demonstrated that the possible whereabouts are bound by an ellipse, with foci at the respective point-locations of the consecutive samples. However, the first works to present a formal analysis of the properties of the so called *beads* are [7]. An illustration is provided in Figure 1.

The corresponding bead (or, space-time prism) is specified with the following constraints:

$$\begin{cases} t_i \leq t \leq t_{i+1} \\ (x - x_i)^2 + (y - y_i)^2 \leq [(t - t_i)v_{max}^i]^2 \\ (x - x_{i+1})^2 + (y - y_{i+1})^2 \leq [(t_{i+1} - t)v_{max}^i]^2 \end{cases} \quad (1)$$

where  $v_{max}$  is the maximal speed that the object can take between  $t_i$  and  $t_{i+1}$ . As shown in Figure 1, at any time instant  $t$  between two consecutive samples, the possible locations of the objects are bound by the lens – i.e., intersection of two circles centered at the respective foci and with respective radii  $v_{max}(t - t_1)$  and  $v_{max}(t_2 - t)$ .

If the objects are constrained to move along a road network, then the space-time prisms are restricted in their size. Specifically, if the segments of the road network are assumed to be edges in a graph, then the prisms become restricted to planar figures [3].

Throughout this paper, we consider the following definition of a trajectory:

**Definition 1.** A trajectory  $Tr_i$  of a moving object with a unique identifier (*oID*) “ $i$ ”, is a sequence of triplets  $Tr = [(L_1, t_1), (L_2, t_2), v_{max1}] \dots$ ,

$[(L_{n-1}, t_{n-1}), (L_n, t_n), v_{max_{(n-1)}}]$  where each  $L_i = (x_i, y_i)$  is a point in 2D space in a corresponding reference coordinate system, and  $t_i$  denotes the time instant at which the object was at location  $L_i$ . When it comes to the time-values,  $i < j$  implies  $t_i < t_j$ , and  $v_{max_{i}}$  denotes the maximum speed of the object between samples at  $t_i$  and  $t_{i+1}$

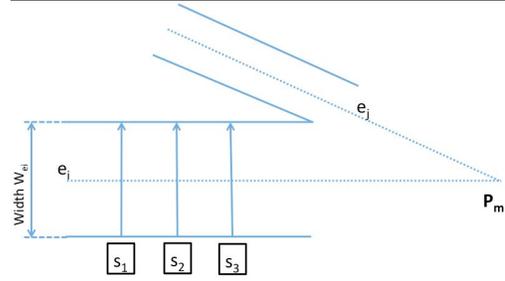


Figure 2: Road Segments and Sensors

We define a road network as an *augmented graph*  $G = (P, E_{RS})$  where  $P = \{p_1, p_2, \dots, p_n\}$  denotes a set of points (commonly corresponding to intersections) and  $E_{RS} = \{r_{S1}, \dots, r_{Sk}\}$  is a collection of triplets of the form  $r_{Si} = (e_i, w_{ei}, v_{ei})$  where:

- $e_i = (p_{i1}, p_{i2}) (\in P \times P)$  is a “regular edge” (i.e., a link between two connected vertices)
- $w_{ei}$  denotes the width of the road segment associated with the edge  $e_i$ .
- $v_{ei}$  denotes the maximum speed associated with  $r_{Si}$ .

We assume that the maximum speed in-between two consecutive location samples along a particular road segment corresponds to the speed-limit of that segment. Geometrically speaking, the collection of all the  $r_{Si}$ 's is the boundary of the Minkowski sum of each “regular edge”  $e_i$  and a disk with diameter  $w_{ei}$ .

We also assume the existence of a collection of sensors  $S = \{s_1, s_2, \dots, s_m\}$ , where each sensor  $s_j$  is located at a point along the outer boundary of some road segment  $r_{Si}$ . Each  $s_j$  detects when (i.e., the time instant at which) a moving object crosses the line segment going through its location and perpendicular to  $e_i$ . The concepts are illustrated in Figure 2.

### 3. FUSING UNCERTAINTIES

We now introduce the new uncertainty model resulting from combining the GPS-based location data and the location data generated by road-side sensors.

The main observation is that the road-side sensors provide additional constraints on the possible whereabouts in-between two consecutive GPS-based samples (and vice-versa). In addition to the system of inequalities (1) specifying the bead, we now have the constraint that at a particular time instant  $t_{si}$ , the locations of the objects are known to also be along a given line-segment determined by: (1) the location of the corresponding road-side sensor; and (2) the direction which is perpendicular to the (boundaries of the) road segment, formalized as:

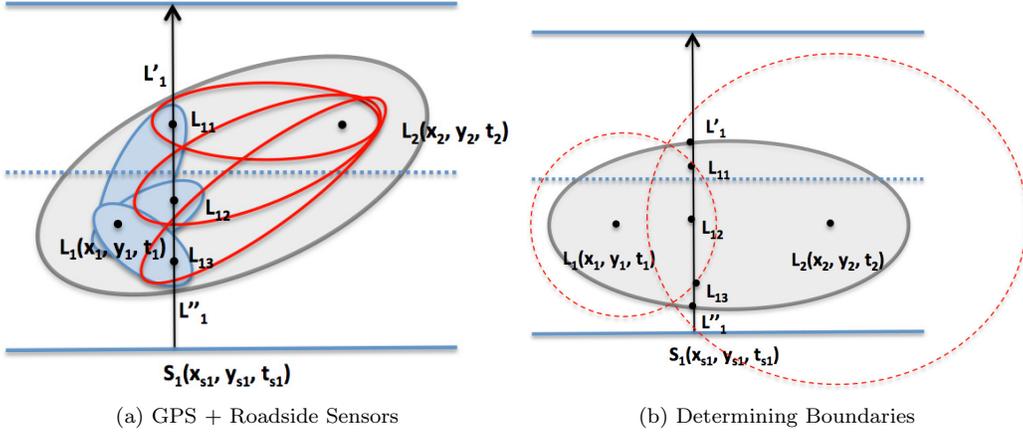


Figure 3: Fusing GPS and Roadside Sensors Data

$$\begin{cases} t_i \leq t \leq t_{i+1}, \\ (x - x_i)^2 + (y - y_i)^2 \leq (t - t_i)^2 v_{max}^2, \\ (x - x_{i+1})^2 + (y - y_{i+1})^2 \leq (t_{i+1} - t)^2 v_{max}^2, \\ y = m_i x + b_i, \text{ when } t = t_{si} \\ t_i \leq t_{si} \leq t_{i+1}. \end{cases} \quad (2)$$

The system of constraints (2) is illustrated in Figure 3: the original GPS-based locations  $L_1$  and  $L_2$  would yield a 2D projection which is an ellipse having them as foci (light-grey shade in Figure 3a) – denote it  $El_1$ . Due to the road-side sensor, the possible locations of the moving object at  $t_{s1}$  can only be along the *portion* of line segment originating in  $(x_{s1}, y_{s1})$ , perpendicular to the boundaries of the road segment, and intersecting  $El_1$  – i.e., along the portion of the line segment  $L'_1 L''_1$ . Clearly, that intersection has an uncountably many points, and we show 3 such points in Figure 3a –  $L_{11}, L_{12}$  and  $L_{13}$ . Each such point, in turn, can be used as a “generator” for two more space-time prisms: one originating in  $L_1$ , and the other terminating at  $L_2$ . The corresponding 2D projections (ellipses) are shown in Figure 3a for  $L_{11}, L_{12}$  and  $L_{13}$ . The most important implication is that when combining the original ellipse  $El_1$  with the uncountably infinite collection of the ellipses with one of the foci along the line segment due to the road-side sensors, the additional constraint induces a significant amount of a “dead-space” in  $El_1$ . A more detailed illustration of the valid range for selecting the points that will generate the infinite collection of (pairs of) new beads is given in Figure 3b. Recall that at any  $t_{s1}$  between the sampling times  $t_1$  and  $t_2$ , the object can be located inside of the lens obtained as the intersection of the circles with radii  $v_{max}(t_{s1} - t_1)$  and  $v_{max}(t_2 - t_{s1})$ . Hence, although the ray emanating from the roadside sensor  $s_1$  would intersect the “global boundary” (i.e., the ellipse which is the projection of the bead) at  $L'_1$  and  $L''_1$ , the only valid points to be considered as possible whereabouts are the ones along (and inside) the lens. As shown in Figure 3b, those are the points along the line segment bounded by  $L_{11}$  and  $L_{13}$ .

We note that there is the “flip-side” context of having a single uncertainty source: if we only had the roadside sensors available, then, in between two detections by consecutive sensors (say,  $s_1$  and  $s_2$  from Figure 2), the whereabouts

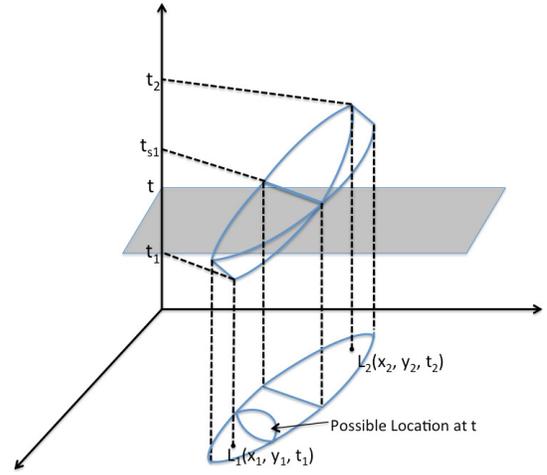


Figure 4: Whereabouts at Time Instant

of a given object is bounded by the union  $\cup(El_{si,sj})$  of uncountably many ellipses for which:

1. The first focus is some point  $L_{s1}$  located on the line-segment originating at the location of  $s_1$ .
2. The second focus is some point  $L_{s2}$  located on the line-segment originating at the location of  $s_2$ ;
3. The distance between  $L_{s1}$  and  $L_{s2}$  is smaller than  $v_{max}(t_{s2} - t_{s1})$  (i.e., the object could travel the distance within the time-interval  $[t_{s1}, t_{s2}]$  for the given speed limit).

The possible whereabouts of the object at a given time instant – i.e., *where\_at(oID, t)* query – for bead obtained by GPS-based samples, is determined via intersecting that bead with a horizontal plane at  $t$  (cf. Figure 1) – i.e., intersection of two circles centered at  $L_1$  and  $L_2$  with the radii corresponding to  $v_{max}(t - t_1)$  and  $v_{max}(t_2 - t)$ .

To determine the whereabouts at a given time instant  $t$  for a fused bead, we need to obtain the intersection of  $FB$  with the horizontal plane  $Time = t$ . The corresponding illustration of the volume in 2D space + Time, along with the 2D projection, is shown in Figure 4. We note that the boundary of the 2D projection is obtained as the “envelope” of the union of two collections of uncountably many infinite pairs of arcs. Each pair of arcs represents the boundaries of

the intersections of the corresponding pairs of disks – one centered at the focus of the GPS-based bead (e.g.,  $L_1$ ) and the other centered at a point along the intersection chord (cf.  $L_{11}L_{13}$  in Figure 3) resulting from secant due to the roadside sensor and the arc from the lens of the GPS-based bead. Thus, one of the boundaries is always a circular arc originating at the focal point of the "original" GPS-based bead, centered at focus of the GPS-based bead (say,  $L_1$ ) and with radius  $v_{max}(t - t_1)$ . The the boundary is actually the boundary of the union of uncountably many disks with radii  $v_{max}(t_s - t)$ , with centers along the intersection-chord.

The complementary query, *when\_at(oID, L)* returns the times during which it is possible for the object *oID* to be at the location  $L(x_L, y_L)$ , i.e., a time-interval  $[t_L^1, t_L^2]$ . The time-interval can be defined as the two intersections between the boundary of the fused bead *FB* and the vertical line (i.e., ray) emanating from *L*. To calculate the values, we have the following observations:

1.  $t_L^2$  is the latest time that a circle located at the GPS-based focus from the sample at  $t_1$  will "reach" *L* – hence, it can be obtained as a solution to the equation:

$$\overline{L_1L} = v_{max}(t_L^2 - t_1)$$

2.  $t_L^1$ , on the other hand, is the earliest time that any circle with the center on the intersection chord and radius  $v_{max}(t_s - t_L^1)$  would pass through *L*.

To get some quantitative evaluation of the proposed model, we examined how many answers obtained using the GPS-based bead model actually become false positive when the *FB* model is employed. Towards that, we used a MATLAB implementation of the numerical integration<sup>1</sup> for evaluating the probabilities of an object satisfying the lane-crossing query. We investigated the impact of *FB* model on lane-crossing query on road networks:

$Q_{lc}^u$ : Retrieve all the vehicles which have  $> \Theta$  ( $0 < \Theta \leq 1$ ) probability of crossing the lane in road segment *RS1*.

We simulated a vehicle moving along a road (cf. Figure 2) and having multiple crossings of the lane, with two direction-components: one parallel to the boundary of the road (i.e., lane), denoted as  $M_x$ , and the other one perpendicular to it,  $M_y$ . The width of each lane of the road segment was set to 4m; the central lane was set as x-axis with full range for  $M_y$  is [-4m, 4m], and densely deployed sensors – located at every 10m along the road. Vehicle's GPS positions were sampled every 1s and the movement along  $M_x$  was set to a constant speed  $< 50$ km/h. Vehicle's perpendicular movement  $M_y$  is generated by a random generator with uniform distribution given a movement interval. Figure 5 shows the benefit of *FB*. The number of lane-crossing incidents is reduced by around 30% by using *FB* model. The reductions are the result of correctly classification for those boundary scenarios that would misclassified as false positive by GPS-based bead model.

#### 4. CONCLUDING REMARKS

We proposed a formal model – fused bead – for the possible locations at given time-instant(s) when the location data from two sources – on-board GPS devices and roadside sensors – is combined, and demonstrated that "two uncertainties are better than one", in the sense that fusing the data from both sources would narrow the possible whereabouts.

<sup>1</sup>The source code(s) and the data are publicly available at <http://www.eecs.northwestern.edu/~bzv686/FusedBeads>

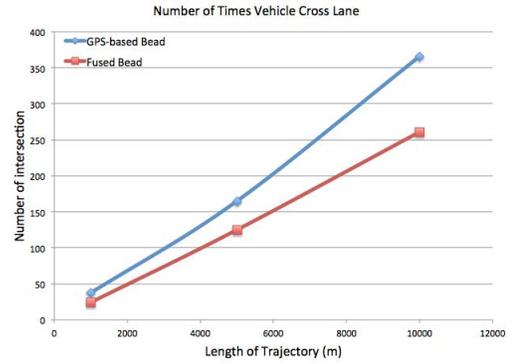


Figure 5: Lane-Crossing Query

We analyzed the impact of the model on the lane-crossing query and we presented experimental observations illustrating the benefits of the fused bead approach.

In the future, we will work on generating efficient algorithms for *continuous* spatio-temporal queries. and we will investigate the scalability and efficiency aspects of the query processing algorithms.

#### 5. REFERENCES

- [1] Southwest research institute, advanced traffic management systems. <http://www.swri.org/4ORG/d10/its/atms/default.htm>.
- [2] T.-F. H. R. Center. *Traffic Detector Handbook*. U.S. Department of transportation.
- [3] T. Emrich, H.-P. Kriegel, N. Mamoulis, M. Renz, and A. Züfle. Querying uncertain spatio-temporal data. In *ICDE*, pages 354–365, 2012.
- [4] R. H. Güting and M. Schneider. *Moving Objects Databases*. Morgan Kaufmann, 2005.
- [5] M. K. G. Institute. Big data: The next frontier for innovation, competition, and productivity, 2011.
- [6] J. Jang, H. Kim, and H. Cho. Smart roadside server for driver assistance and safety warning: Framework and applications. In *CUTE 2010*, pages 1–5, Dec 2010.
- [7] B. Kuijpers and W. Othman. Trajectory databases: data models, uncertainty and complete query languages. *Journal of Computer and System Sciences*, 2009. doi:10.1016/j.jcss.2009.10.002.
- [8] T. Leonard. Delivering deeper insights with big data and real-time analytics, 2012.
- [9] D. Pfoser and C. S. Jensen. Capturing the uncertainty of moving objects representation. In *SSD*, 1999.
- [10] R. Toledo-Moreo, D. Betaille, and F. Peyret. Lane-level integrity provision for navigation and map matching with gnss, dead reckoning, and enhanced maps. *Intelligent Transportation Systems, IEEE Transactions on*, 11(1):100–112, March 2010.
- [11] G. Trajcevski, R. Tamassia, I. Cruz, P. Scheuermann, D. Hartglass, and C. Zamierowski. Ranking continuous nearest neighbors for uncertain trajectories. *VLDB J.*, 20(5):767–791, 2011.
- [12] G. Trajcevski, O. Wolfson, K. Hinrichs, and S. Chamberlain. Managing uncertainty in moving objects databases. *ACM Trans. Database Syst.*, 29(3), 2004.