Manage the Data from Indoor Spaces: Models, Indexes & Query Processing

Huan Li

Database Laboratory, Zhejiang University

lihuancs@zju.edu.cn

March 23, 2016

Overview

About This Work...

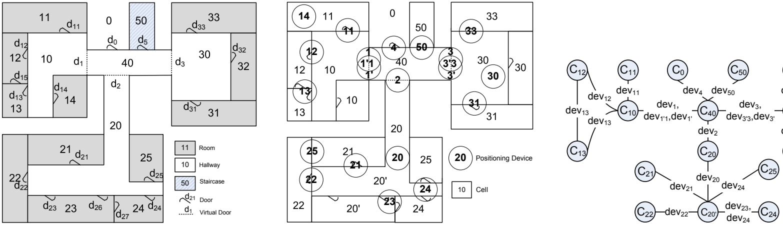
Spatio-temporal Joins on Symbolic Indoor Tracking Data. [4]
H. Lu, B. Yang, and C. S. Jensen.

- Published at *ICDE' 2011*.
- Studies the probabilistic, spatio-temporal joins on hisorical indoor tracking data.
- Two-phase hash-based algorithms are proposed for the point and interval joins.
- A filter-and-refine framework, along with spatial indexes and pruning rules.

Motivation

- Huge amount of tracking data serves as a foundation for a wide variety of indoor applications and services.[1]
  - shopping mall, airports, office buildings, akin to those enabled by outdoor GPS
  - hot area detection, space planning, security control, movement pattern discovery
- Spatio-temporal joins fall short in indoor setting.
  - indoor space consists of semantic entities enable or constrain movement
  - semantics of indoor space call for novel spatio-temporal join predicates
  - indoor positioning technologies differ fundamentally from outdoor setting, low accuracy and sampling frequency
- Joins on indoor tracking data call for new definition and new implementation techniques that take into account:
  - specifics of indoor space
  - limitations of indoor positioning

Preliminaries: Symbolic Indoor Tracking



(a) Floor Plan (b) Devices Deployment (c) Deployment Graph

1. $C2P : C \to 2^P$ maps a cell to a set of indoor partitions
2. $D2C : D \to 2^C$ maps a device to a set of corresponding cells
3. According to Deployment Graph, for partitioning device, $D2C(device_{13}) = \{C_{10}, C_{13}\} \cup \{C_{12}, C_{13}\} = \{C_{10}, C_{12}, C_{13}\}$
4. For presence device, $D2C(device_{25}) = \{C_{21}, C_{22}\}$ as the cells intersect its detection range.
5. $D2C : D \to 2^C$ is useful as it captures the possible movements of objects.

Preliminaries: Symbolic Indoor Tracking

TABLE I
OBJECT TRACKING TABLE (OTT)

ID	$objectID$	$deviceID$	t_s	t_e
rd_1	o_1	dev_4	t_1	t_2
rd_2	o_2	dev_4	t_1	t_2
rd_3	o_1	dev_2	t_5	t_6
rd_4	o_2	$dev_{1'}$	t_7	t_8
rd_5	o_1	dev_1	t_9	t_{10}
rd_6	o_1	dev_{12}	t_{15}	t_{16}
rd_7	o_2	dev_{13}	t_{20}	t_{21}
rd_8	o_1	dev_{13}	t_{21}	t_{22}
rd_9	o_2	dev_{13}	t_{29}	t_{30}
...	...	...	...	...

- **Object Tracking Table** OTT records the converted trajectories with schema $(ID, objectID, deviceID, t_s, t_e)$
- a record states that the object $objectID$ is observed by the device $deviceID$ in the closed interval from time t_s to t_e.

Problem Definitions

Given an OTT, it is of interesting to identify object pairs that join w.r.t some specific spatio-temporal join predicate.

- to know all pair of individuals that were probably at the same gate when a particular event (terrorist attack) occurred in a large airport.

Due to tracking uncertainty, only interested in those objects that satisfy the join predicate with some given probability (specified threshold).

The joins are effectively *self-joins* because all tracking data is contained in a single OTT.

Problem Definition I

One can apply a join predicate to a time point to find pairs that join at that particular time point...

Definition (Probabilistic Threshold Indoor Spatio-temporal Join–PTISSJ)

Given an OTT, a join predicate P, a time point t, and a threshold value $M \in (0, 1]$, a probabilistic threshold indoor spatio-temporal join $\bowtie_{P,t,M} (OTT) = \{(o_i, o_j) | o_i, o_j \in O \wedge o_i \neq o_j \wedge pr(P(o_i, o_j, t)) > M\}$, where $pr(P(o_i, o_j, t))$ is the **Timeslice Join Probability** of o_i, o_j at time t, i.e., the probability that predicate $P(o_i, o_j, t)$ is true.

Problem Definition II

It's also interesting to know object pairs satisfy the predicate for some consecutive timestamp...

Definition (Probabilistic Threshold k Indoor Spatio-temporal Join–PTkISSJ)

Given an OTT, a join predicate P, a time interval $I = [t_m, t_n](m < n)$, an integer $k(0 < k \leq n - m)$, and a threshold value $M \in (0, 1]$, a probabilistic k threshold indoor spatio-temporal join

$$\bowtie_{P,I,k,M} (OTT) = \{(o_i, o_j) | o_i, o_j \in O \wedge o_i \neq o_j \wedge$$
$$\exists s \in m...n - k + 1(\forall \delta \in 0...k - 1(pr(P(o_i, o_j, t_{s+\delta})) > M))\}$$

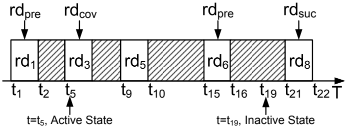
Uncertainty Model for Indoor Tracking

For outdoor moving objects [2], **Uncertainty Region**, denoted by $UR(o_i, t)$, is a region such that o_i must be in this region at time t.

In general terms, an object o_i's location can be modeled as a random variable l associated with a probability density function $f_{o_i}(l, t)$ that has non-zero values only in o_i's suncertainty region $UR(o_i, t)$. [3]

$$\int_{l \in UR(o_i,t)} f_{o_i}(l, t) dl = 1 \tag{1}$$

Object State in OTT



Definition (Active State)

Given an object o_i and a time point t, if a tracking record rd_{cov} is found in OTT such that $rd_{cov}.objectID = o_i$ and $t \in [rd_{cov}.t_s, rd_{cov}.t_e]$, o_i is in the **active state** at time t.

Definition (Inactive State)

Given an object o_i and a time point t, if no record rd_{cov} is found in OTT, o_i is in the **inactive state** at time t. Instead, two tracking records of o_i called rd_{pre} and rd_{suc}, can be found in OTT, such that they are consecutive in the sense that $rd_{pre}.t_e < t < rd_{suc}.t_s$ and there is no record for o_i between times $rd_{pre}.t_e$ and $rd_{suc}.t_s$.

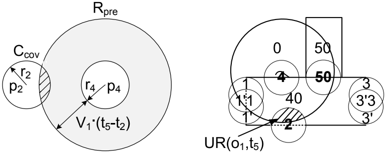
Uncertainty Region in the Active State

TABLE I
OBJECT TRACKING TABLE (OTT)

ID	$objectID$	$deviceID$	t_s	t_e
rd_1	o_1	dev_4	t_1	t_2
rd_2	o_2	dev_4	t_1	t_2
rd_3	o_1	dev_2	t_5	t_6
rd_4	o_2	$dev_{1'}$	t_7	t_8
rd_5	o_1	dev_1	t_9	t_{10}
rd_6	o_1	dev_{12}	t_{15}	t_{16}
rd_7	o_2	dev_{13}	t_{20}	t_{21}
rd_8	o_1	dev_{13}	t_{21}	t_{22}
rd_9	o_2	dev_{13}	t_{29}	t_{30}
...	...	...	...	...

Example

$t = t_5$, $rd_{cov} = rd_3$ and $rd_{pre} = rd_1$, which tells o_i left dev_4's detection range at time t_2, and is currently detected by dev_2.



Step 1: UR is the detection range of device $rd_{cov}.deviceID$, denote as:

$$C_{cov} = Cir(Loc(rd_{cov}.deviceID), Rad(rd_{cov}.deviceID))$$

Step 2: UR should consider the rd_{pre}'s *maximum speed bounding ring*(MSBR):

$$UR(o_i, t) = C_{cov} \cap Ring(Loc(rd_{pre}.deviceID), Rad(rd_{pre}.deviceID), V_i \cdot (t - rd_{pre}.t_e))$$

Uncertainty Region in the Inactive State

TABLE I
OBJECT TRACKING TABLE (OTT)

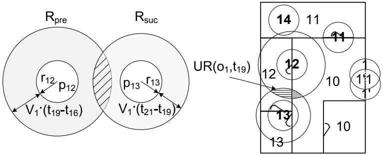
ID	$objectID$	$deviceID$	t_s	t_e
rd_1	o_1	dev_4	t_1	t_2
rd_2	o_2	dev_4	t_1	t_2
rd_3	o_1	dev_2	t_5	t_6
rd_4	o_2	$dev_{1'}$	t_7	t_8
rd_5	o_1	dev_1	t_9	t_{10}
rd_6	o_1	dev_{12}	t_{15}	t_{16}
rd_7	o_2	dev_{13}	t_{20}	t_{21}
rd_8	o_1	dev_{13}	t_{21}	t_{22}
rd_9	o_2	dev_{13}	t_{29}	t_{30}
...	...	...	...	...

Example

$t = t_{19}$, $rd_{pre} = rd_6$ and $rd_{suc} = rd_8$,

since $rd_6.t_e = t_{16} < t_{19} < rd_8.t_s = t_{21}$.

we have $dev_p = dev_{12}$ and $dev_s = dev_{13}$



Step 1: Determine the possible cells in which the object can be in the inactive period:

$$Cells_{mid} = D2C(dev_p) \cup D2C(dev_s)$$

Step 2: UR is constrained by two *maximum speed bounding ring*(MSBR)s of rd_{pre} and rd_{suc}:

$$UR(o_i, t) = \bigcup_{c \in Cells_{mid}} c \cap R_{pre} \cap R_{suc}$$

References

C. S. Jensen, H. Lu, and B. Yang.
Graph model based indoor tracking.
In *MDM*, pp. 122–131, 2009.

B. Yang, H. Lu, and C. S. Jensen.
Scalable continuous range monitoring of moving objects in symbolic indoor space.
In *CIKM*, pp. 671–680, 2009.

B. Yang, H. Lu, and C. S. Jensen.
Probabilistic threshold k nearest neighbor queries over moving objects in symbolic indoor space.
In *EDBT*, pp. 335–346, 2010.

H. Lu, B. Yang, and C. S. Jensen.
Spatio-temporal Joins on Symbolic Indoor Tracking Data.
In *ICDE*, pp. 816–827, 2011.

References

C. S. Jensen, H. Lu and B. Yang.
Indoor-A New Data Management Frontier.
In *IEEE Data Eng. Bull.*, pp. 12–17, 2010.

Reynold Cheng, Dmitri V Kalashnikov and Sunil Prabhakar.
Querying imprecise data in moving object environments.
In *TKDE*, pp. 1112–1127, 2004.

The End. Thanks :)