Prediction of Moving Object Location Based on Frequent Trajectories

Mikołaj Morzy

Institute of Computing Science
Poznań University of Technology
Piotrowo 2, 60-965 Poznań, Poland

The 21st International Symposium on
Computer and Information Sciences ISCIS'2006
Istanbul, Turkey, November 2006

Outline

1 Introduction

2 Related Work

3 Definitions

4 Prediction of Location

5 Experiments

6 Conclusions

Motivation

Observations

- ubiquitous mobile devices
  - mobile phones, PDAs, vehicles
  - GPRS, Bluetooth, Wi-Fi, WiMAX
- advent of location-based services
  - traffic management
  - way-finding
  - location-based advertising
  - location-based information retrieval
- exact position of a moving object rarely known
  - periodicity of position disclosure
  - existence of urban canyons
  - natural phenomena
  - power shortages

Current solutions

- Complex models using network topology
  - yield accurate results, but computationally unfeasible
- Simulation-based models
  - numerous parameters governing the model
  - cost of computation may be prohibitively high
  - adaptation to dynamic changes in environment

Current solutions

- Complex models using network topology
  - yield accurate results, but computationally unfeasible
- Simulation-based models
  - numerous parameters governing the model
  - cost of computation may be prohibitively high
  - adaptation to dynamic changes in environment

Problem

Current techniques make little or no use of the huge amounts of historical data generated by moving objects

Mining mobile object data

- Prediction accuracy vs. prediction speed
- Movement data acquired from moving objects hide valuable knowledge about moving object behavior, but ...

Mining mobile object data

- Prediction accuracy vs. prediction speed
- Movement data acquired from moving objects hide valuable knowledge about moving object behavior, but ...

Question

Are data mining techniques too slow and too computationally expensive for real-time location prediction?

Requirements and thesis

A method for location prediction must:

- produce reliable predictions
- explain predictions
- perform in near real-time
- utilize historical data

Requirements and thesis

A method for location prediction must:

- produce reliable predictions
- explain predictions
- perform in near real-time
- utilize historical data

Thesis

Data mining techniques are appropriate and efficient in real-time location prediction under assumption that enough historical data exist

Solution and contribution

Our solution

1. superimpose a grid on the movement area
2. transform movement paths into trajectories expressed in terms of grid edges
3. discover frequent trajectories
4. transform frequent trajectories into movement rules
5. match the history of an object with the database of movement rules
6. produce a probabilistic model of possible object locations

Solution and contribution

Our solution

1. superimpose a grid on the movement area
2. transform movement paths into trajectories expressed in terms of grid edges
3. discover frequent trajectories
4. transform frequent trajectories into movement rules
5. match the history of an object with the database of movement rules
6. produce a probabilistic model of possible object locations

Our contribution

- using historical data to build environment model
- designing the *AprioriTraj* mining algorithm
- development of four matching strategies
- experimental evaluation of the proposal

Related work

Tracking of moving objects

- Y. Tao, C. Faloutsos, D. Papadias, and B. Liu. Prediction and indexing of moving objects with unknown motion patterns. In *ACM SIGMOD'04, Paris, France, June 13-18*, pp. 611–622. ACM, 2004.
- B. Xu and O. Wolfson. Time-series prediction with applications to traffic and moving objects databases. In *MobiDE 2003, September 19, 2003, San Diego, California, USA*, pp. 56–60. ACM, 2003.
- O. Wolfson and H. Yin. Accuracy and resource concumption in tracking and location prediction. In *SSTD'03, Santorini, Greece, July 24-27*, pp. 325–343. Springer, 2003.
- G. Trajcevski, O. Wolfson, B. Xu, and P. Nelson. Real-time traffic updates in moving objects databases. In *DEXA'02, Aix-en-Provence, France, 2-6 September*, pp. 698–704. IEEE Computer Society, 2002.

Spatio-temporal data mining

- K. Koperski and J. Han. Discovery of spatial association rules in geographic databases. In *SSD'95, Portland, Maine, USA, August 6-9*, pp 47–66. Springer, 1995.
- M. Ester, A. Frommelt, H.-P. Kriegel, and J. Sander. Spatial data mining: Database primitives, algorithms and efficient dbms support. *Data Mininig and Knowledge Discovery*, 4(2/3):193–216, 2000.
- N. Mamoulis, H. Cao, G. Kollios, M. Hadjieleftheriou, Y. Tao, and D. W. Cheung. Mining, indexing, and querying historical spatiotemporal data. In *ACM SIGKDD'04, Seattle, Washington, USA, August 22-25*, pp 236–245. ACM, 2004.

Mining trajectories of moving objects

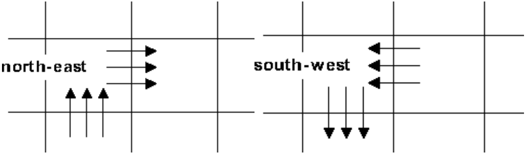
- J. Yang and M. Hu. Trajpattern: Mining sequential patterns from imprecise trajectories of mobile objects. In *EDBT'06, Munich, Germany, March 26-31*, pp 664–681. Springer, 2006.
- Y. Li, J. Han, and J. Yang. Clustering moving objects. In *ACM SIGKDD'04, Seattle, Washington, USA, August 22-25*, pp 617–622. ACM, 2004.

Basic notions

Given a database of moving objects locations, let

- $l_j^i = \left(x_j^i, y_j^i\right)$ denote the i-th location of the j-th object
- $t_j = \left\langle l_j^0, l_j^1, \ldots, l_j^n \right\rangle$ denote the trajectory of the j-th object

Movement area is covered by a grid with cells of constant size, denoted *grid_size*. Each edge, denoted e_{pq}, can be traversed in two directions as follows



Basic notions

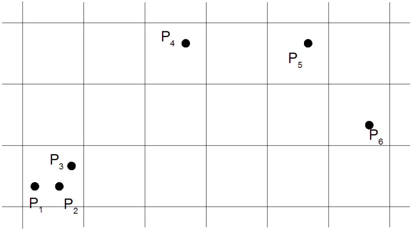
Edges allow to move to a coarser level of granularity

- the trajectory of the j-th object is $t_j = \left\langle (e_{p_0 q_0}, d_0)_j, (e_{p_1 q_1}, d_1)_j, \ldots \right\rangle$, where $d_i \in \{ne, sw\}$ denotes the direction of edge traversal

Basic notions

Edges allow to move to a coarser level of granularity

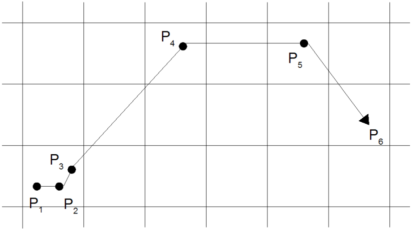
- the trajectory of the j-th object is $t_j = \left\langle (e_{p_0 q_0}, d_0)_j, (e_{p_1 q_1}, d_1)_j, \ldots \right\rangle$, where $d_i \in \{ne, sw\}$ denotes the direction of edge traversal



Basic notions

Edges allow to move to a coarser level of granularity

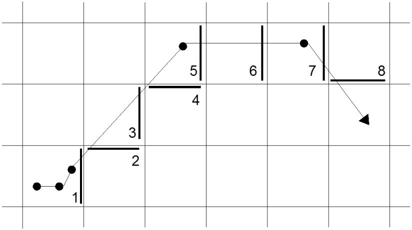
- the trajectory of the j-th object is $t_j = \left\langle (e_{p_0q_0}, d_0)_j, (e_{p_1q_1}, d_1)_j, \ldots \right\rangle$, where $d_i \in \{ne, sw\}$ denotes the direction of edge traversal



Basic notions

Edges allow to move to a coarser level of granularity

- the trajectory of the j-th object is $t_j = \left\langle (e_{p_0 q_0}, d_0)_j, (e_{p_1 q_1}, d_1)_j, \ldots \right\rangle$, where $d_i \in \{ne, sw\}$ denotes the direction of edge traversal



Additional issues

Issues

- support of an edge
- frequent edge
- sub-trajectory of a trajectory
- adjacency of trajectories
- concatenation of trajectories
- apriori properties of frequent trajectories

Movement rules

Definition

An expression of the form $t_i \Rightarrow t_j$ where $t_i, t_j \in L$, t_i and t_j are adjacent trajectories and $t_i \| t_j$ is a frequent trajectory

Properties

$$support\left(t_i \Rightarrow t_j\right) = \frac{\left|t_k \in D : t_k \supseteq \left(t_i \| t_j\right)\right|}{|D|}$$

$$confidence\left(t_i \Rightarrow t_j\right) = P\left(t_j | t_i\right) = \frac{support\left(t_i \| t_j\right)}{support\left(t_i\right)}$$

AprioriTraj algorithm

A modification of the well-known *Apriori* algorithm

- find frequent 1-trajectories
- create candidate 2-trajectories from adjacent frequent 1-trajectories
- iteratively build candidate k-trajectories from overlapping frequent $(k-1)$-trajectories
- no false candidates (!)

AprioriTraj algorithm

A modification of the well-known *Apriori* algorithm

- find frequent 1-trajectories
- create candidate 2-trajectories from adjacent frequent 1-trajectories
- iteratively build candidate k-trajectories from overlapping frequent $(k-1)$-trajectories
- no false candidates (!)

Movement rules generation is straightforward

- split every frequent m-trajectory t_i into $(m-1)$ pairs (t_i', t_i'')
- output rule $t_i' \Rightarrow t_j''$ if $confidence\,(t_i' \Rightarrow t_i'') = support\,(t_i)\,/support\,(t_i') \geq minconf$

Matching strategies

Simple strategy

$$\arg\max_{t_i \Rightarrow t_j} \frac{|t_i|}{|t_q|} * confidence\,(t_i \Rightarrow t_j)$$

- does not consider the length of the consequent
- treats the length of the antecedent linearly

Matching strategies

Polynomial strategy

$$\arg\max_{t_i \Rightarrow t_j} \frac{1}{2}\left(\sqrt{\frac{|t_i|}{|c_1|}} + \sqrt{\frac{|t_j|}{|c_2|}}\right) * confidence\left(t_i \Rightarrow t_j\right)$$

- fair and balanced
- compromise between simple and logarithmic strategies

Matching strategies

Logarithmic strategy

$$\arg\max_{t_i \Rightarrow t_j} \left(w_1 + w_2 * \log_{|c_1|} |t_i| + w_3 * \log_{|c_2|} |t_j| \right) * confidence\,(t_i \Rightarrow t_j)$$

- weights shift emphasis on the confidence factor, the relative length of the antecedent, or the relative length of the consequent of the movement rule

Matching strategies

Aggregate strategy

$$\arg\max_{t_i \Rightarrow t_j} \frac{|t_i|}{|t_q|} * \frac{|t_j|}{|c_2|} * \frac{\sum_{t_x \Rightarrow t_y \in G} |t_y| * confidence(t_x \Rightarrow t_y)}{\sum_{t_x \Rightarrow t_y \in G} |t_y|}$$

- allows similar movement rules to be clustered
- considers the coverage factor, the relative length of the antecedent, and the predictive power of the group of related movement rules
- computationally expensive

Experiments

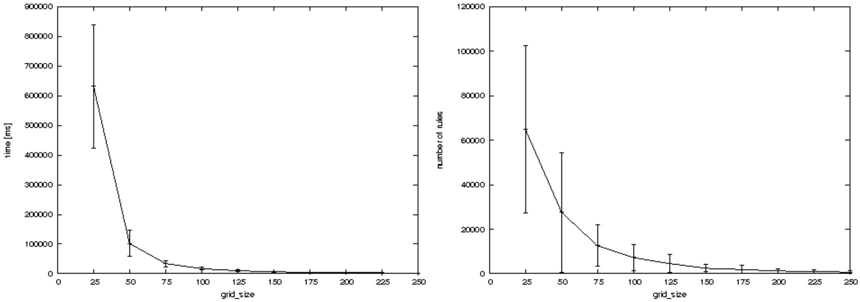
Synthetic data: Network-based Generator of Moving Objects

T.Brinkhoff, A framework for generating network-based moving objects, GeoInformatica, 6(2):153–180, 2002

- maximum velocity: 150
- number of time units: 100
- results averaged over 30 different instances

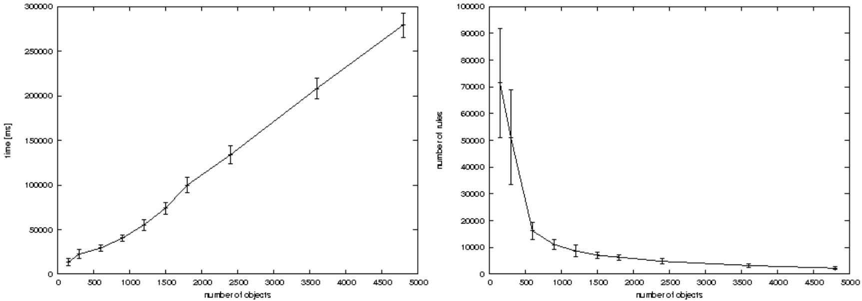
All experiments were conducted on Pentium Centrino 1.8 GHz with 1GB RAM under Windows XP Professional.

Experiment 1



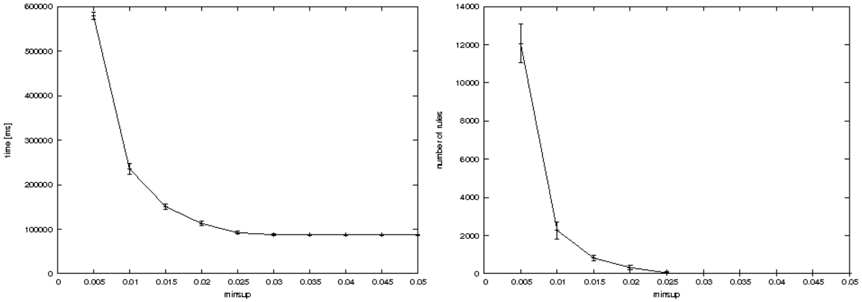
Time and number of rules w.r.t. *grid_size*
300 objects, *minsup* $= 0.03$

Experiment 2



Time and number of rules w.r.t. *number of objects*
grid_size $= 250$, *minsup* $= 0.01$

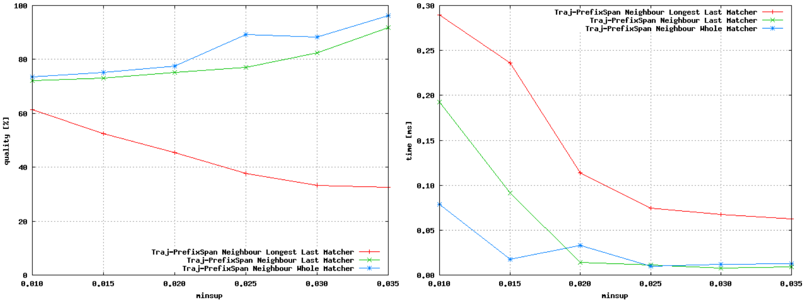
Experiment 3



Time and number of rules w.r.t. *minsup*

$grid\_size = 250$, 4800 objects

Experiment 4



Quality and matching time w.r.t. *minsup*

Conclusions and future work

Conclusions

Movement rules

- provide simplification and generalization of movement patterns
- allow predicting the location of a moving object

Future work

Our future work agenda includes

- comparison of matching strategies
- incorporating temporal aspects
- combining movement rules with spatial data