



Location Order Recovery for Trails with Low Time Resolution



Binxuan Huang
binxuanh@cs.cmu.edu

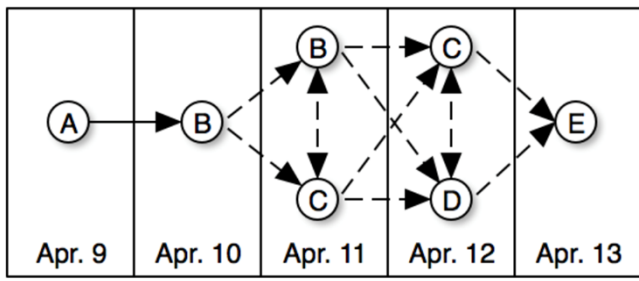
Prof. Kathleen M. Carley
kathleen.carley@cs.cmu.edu

Motivation

Trail: a path of an object through time and space
High time resolution vs. Low time resolution

Time	Location	Time	Location
2016/4/9 10:00	A	2016/4/9	A
2016/4/10 11:00	B	2016/4/10	B
2016/4/11 9:30	B	2016/4/11	B
2016/4/11 13:00	C	2016/4/11	C
2016/4/12 10:00	C	2016/4/12	C
2016/4/12 11:30	D	2016/4/12	D
2016/4/13 14:00	E	2016/4/13	E

Broken point: If there are more than 2 different locations appear at one time slot.



Method

Step 1: Extract transition probability $P(A|B) = \frac{P(A,B)}{P(B)}$ from unbroken trail subsequence.

Step 2: Find the location order with highest probability product (NP hard).

$$P(L_1, \dots, L_n) = P(L_1)P(L_2|L_1) \dots P(L_n|L_{n-1})$$

---Exact algorithm

---Greedy from the start node

---Global Greedy: find an transition edge with highest transition probability first, then traverse to the source and target locations.

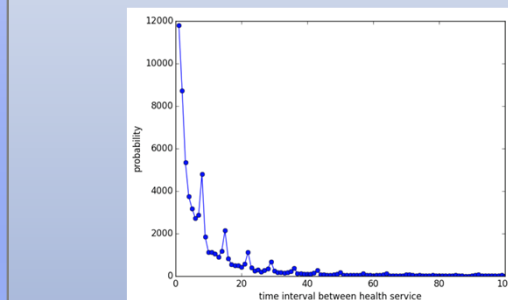
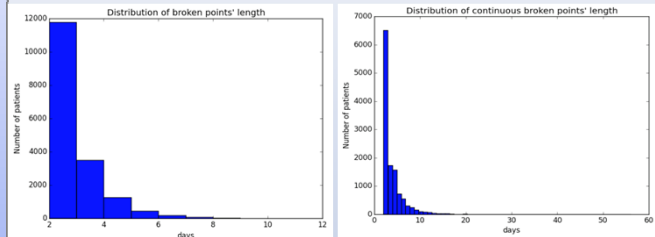
Time interval factor: partition trails when time interval between two locations is larger than a threshold.

Begin/End location: add "BEGIN" node and "END" node at the beginning and end of each partitioned trails.

Dataset

Health service record. It contains individual visits for patients who received care service in a hospital associated Private and Clinical system between July 1 2011 and June 30 2012.

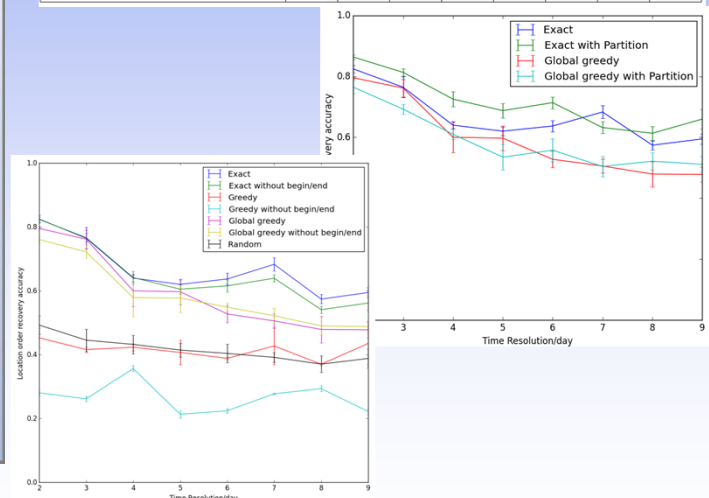
Health data	# of records	# of agents	# of locations	# of unbroken trails
	94885	5055	115	814



Experiments

manually change the time resolution for those unbroken trails: $\text{timestamp}' = \lfloor \text{timestamp} \rfloor * \text{resolution}$.

Time resolution(day)	2	3	4	5	6	7	8	9
# of broken trails	164	213	250	263	268	278	294	297
# of broken points	178	251	296	319	340	353	374	389
Avg. length of broken points	2.073	2.223	2.284	2.426	2.482	2.586	2.591	2.627
Avg. length of cont. broken points	2.121	2.364	2.449	2.651	2.749	2.917	2.910	3.042
Max. length of broken points	4	5	6	5	6	6	6	7



This work was supported in part by the Office of Naval Research (ONR) N000141512563, and the Center for Computational Analysis of Social and Organization Systems (CASOS). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Office of Naval Research or the U.S. government.