Extracting Visited Points of Interest from Vehicle Trajectories

Ilkcan Keles[1], Matthias Schubert[2], Peer Kröger[2], Simonas Šaltenis[1], Christian S. Jensen[1]

[1]Aalborg University, Denmark - {ilkcan, simas, csj}@cs.aau.dk

[2]Ludwig-Maximilians Universität München, Germany - {schubert, kroeger}@dbs.ifi.lmu.de

Introduction

- Motivation
  - The visits extracted from GPS data can be used to evaluate the functionality of location based services.
- Problem
  - Given
    - S_{TR} – A set of GPS trajectories
    - S_P – A database of PoIs in the geographical region covered by S_{TR}
  - Identify the visits of users whose trajectories are given in S_{TR} to the PoIs contained in S_P
- Challenge
  - Finding the actual visited PoI when the stop location is surrounded with many PoIs as shown in the figure

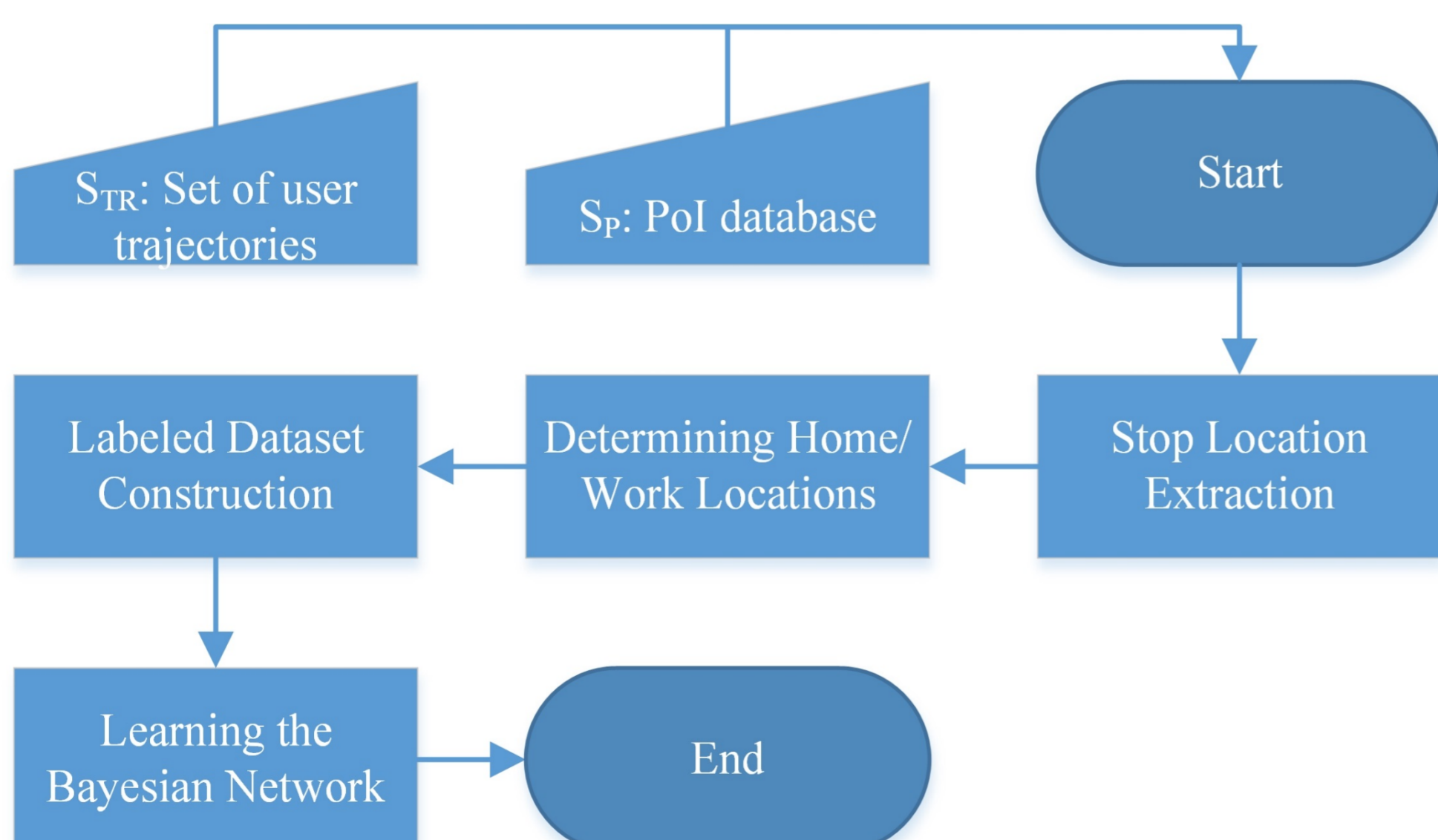- Blue markers: PoIs
- Red markers: Stop Location
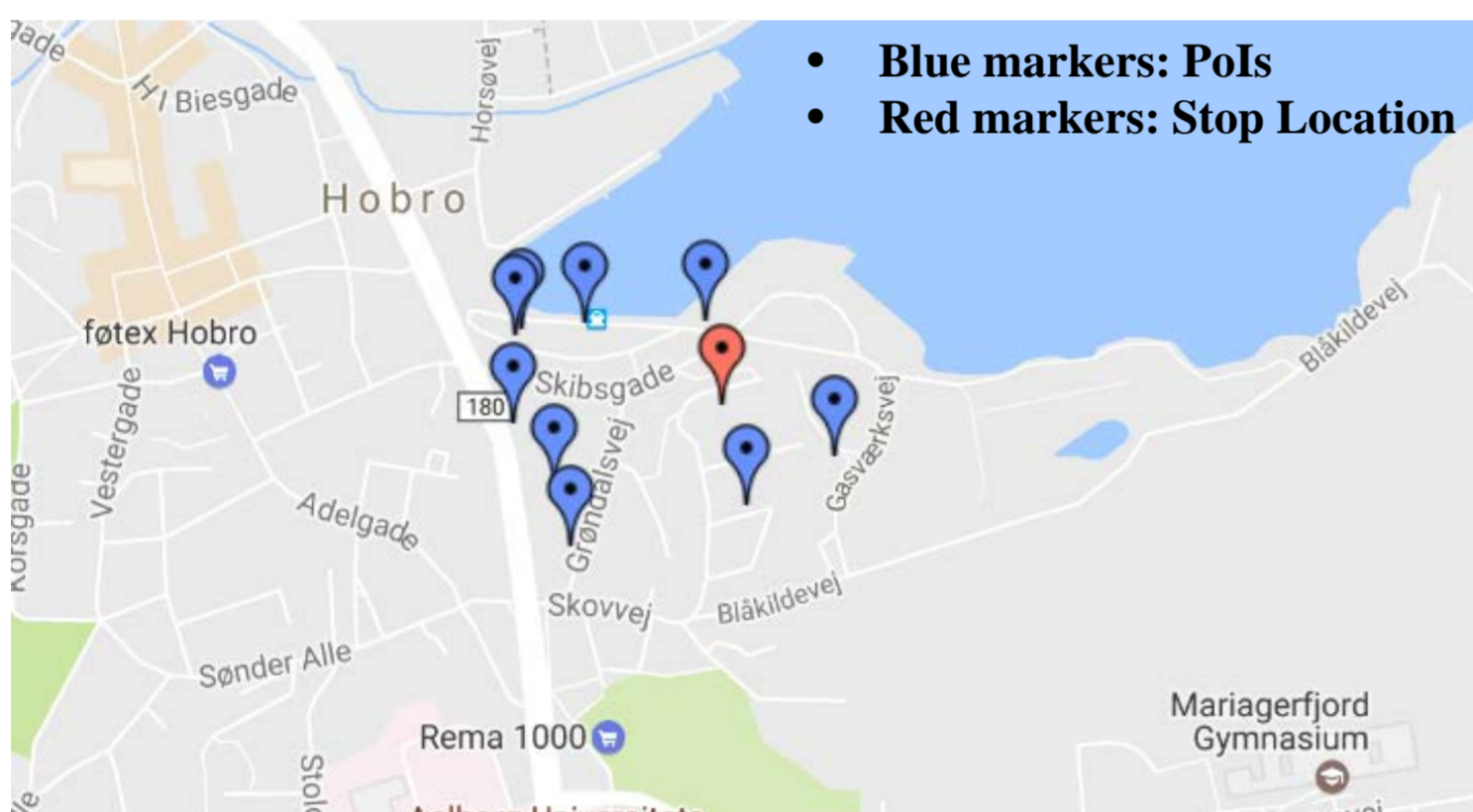
Visited PoI Extraction (VPE) Method

- Bayesian Network Building Phase
  - Construction of labeled assignment data
  - Learning the network
- Assignment Phase
  - The set of possible categories is determined.
  - Joint probability computation using the Bayesian network
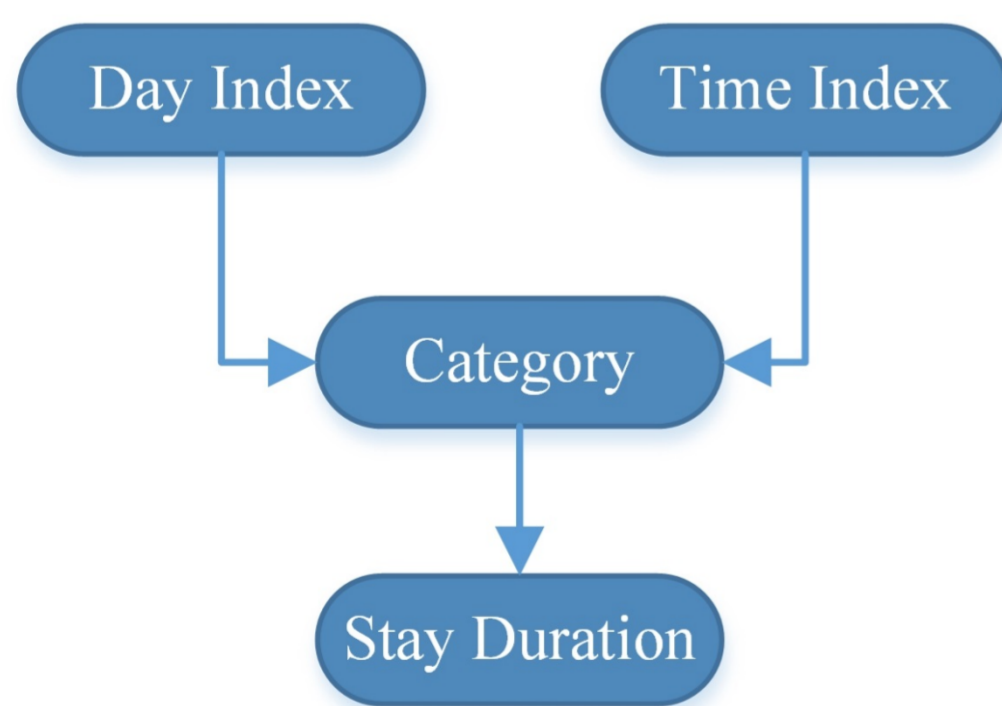

## Building the Bayesian Network

Assignment Phase

- Distance-based filtering to determine the set of possible categories
- Joint probability computation using the Bayesian network
- The category with the maximum probability is determined.
- If there is only one PoI of this category in the set of possible PoIs, the stop location is assigned to this PoI.
  - Otherwise, the stop is not assigned.

Experimental Evaluation

- GPS data containing around 0.4 billion records and 10,000 PoIs of 88 categories are used.
- Ground truth dataset is constructed.
- The method returns a ranked list of categories instead of a single category for evaluation purposes.
- VPE achieves a p@3 value around 0.8 and a mean reciprocal value around 2.
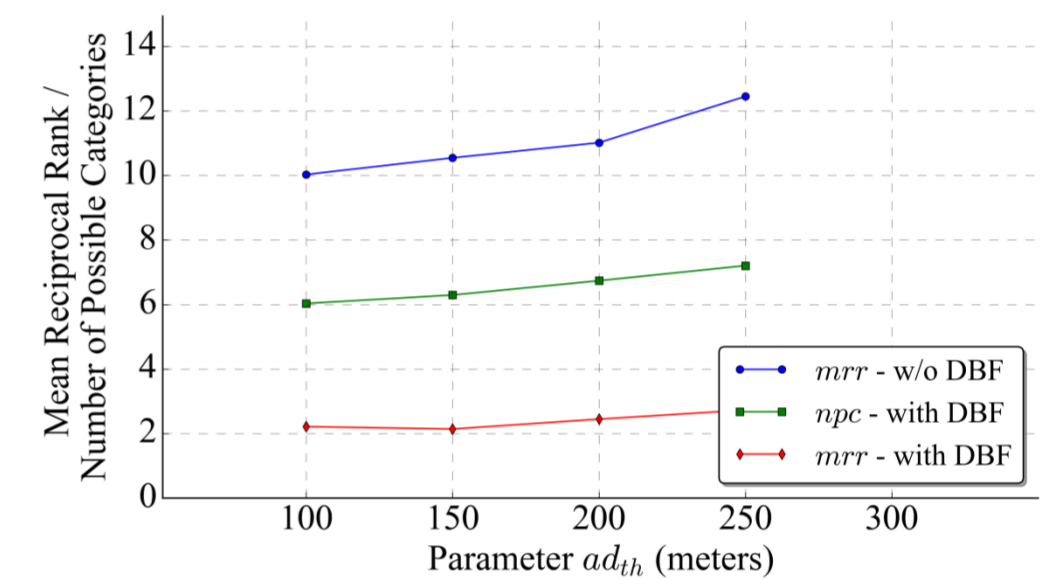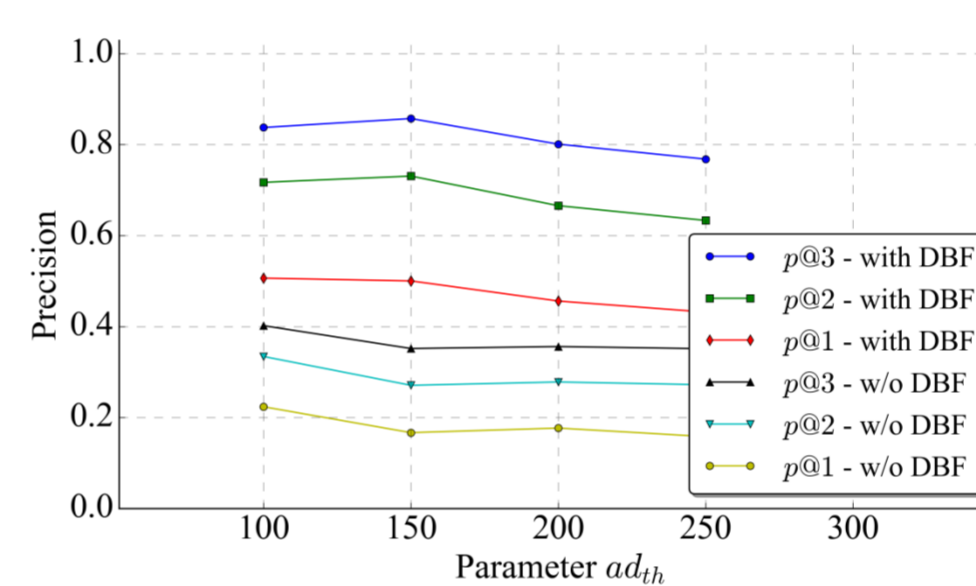
- The best performance is achieved when a time slot is 30 minutes or 1 hour.
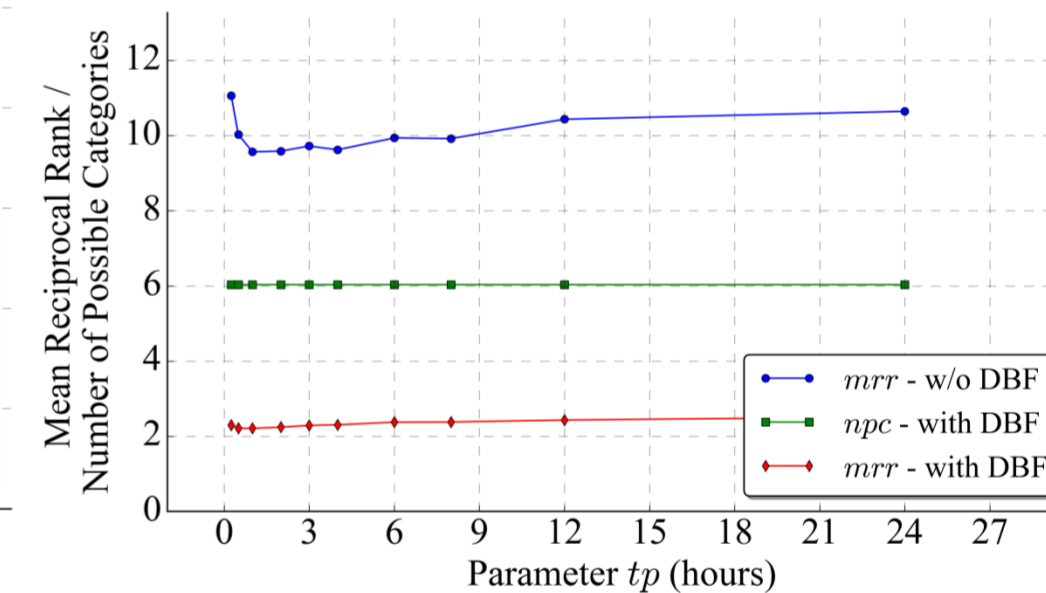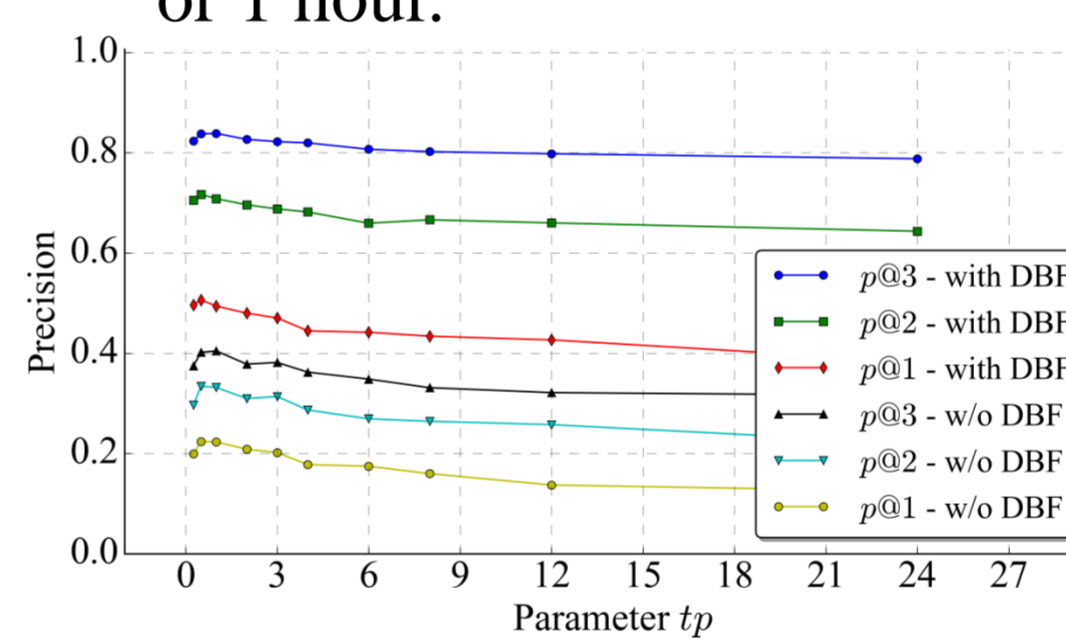
- The stay duration distribution obtained with assignments (right) is quite similar to the one obtained from labeled dataset (left).
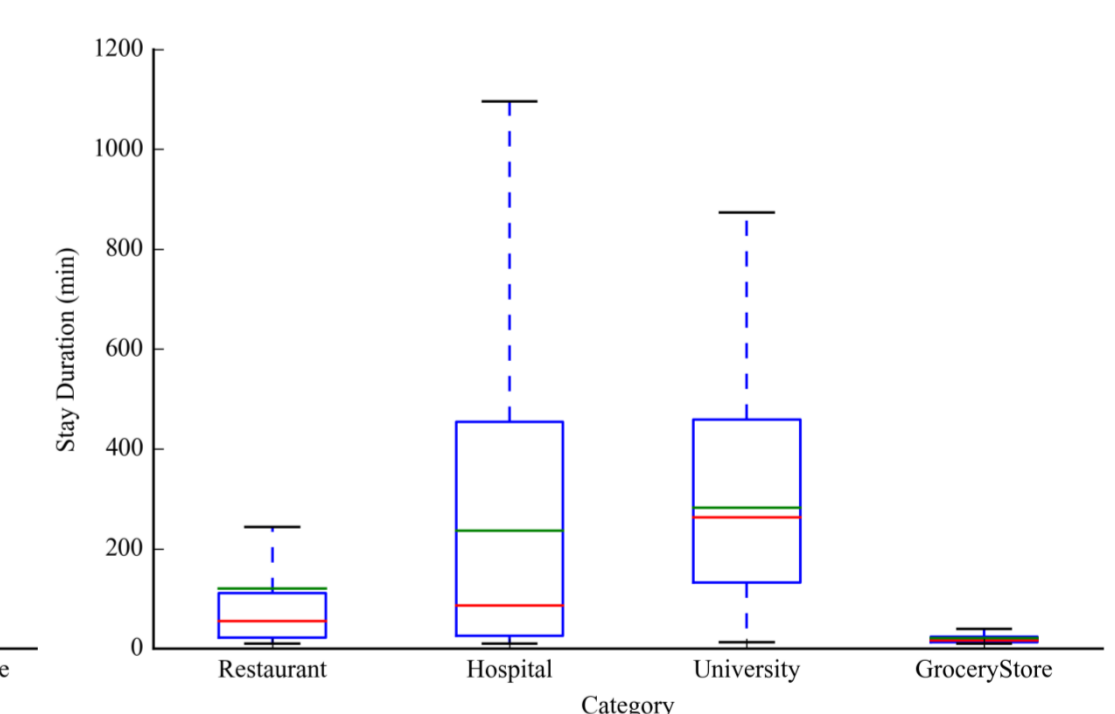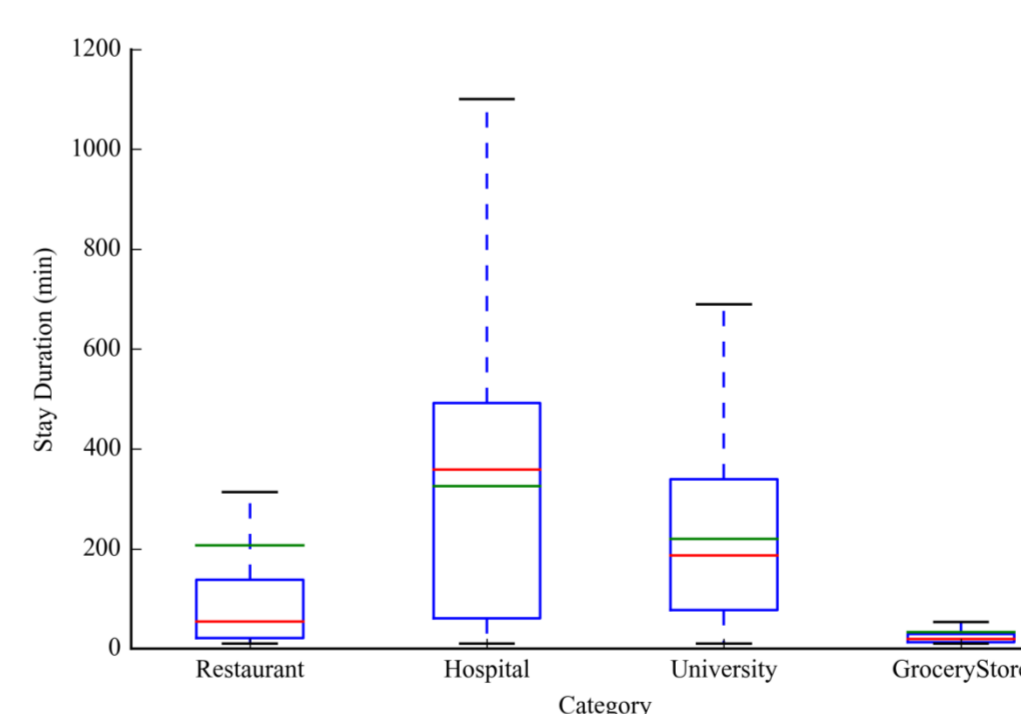
Conclusion

- We propose a Visited PoI Extraction method that
  - Employs a Bayesian network to represent the relationship between the temporal attributes of a stop and the category of the visited PoI
  - Includes a method to build a labeled dataset
- The proposed method is capable of detecting the category of the visited PoI and it achieves a p@3 value of 0.8.

Acknowledgments