Finding Communities of Commuters

Ashish Dandekar, School of Computing, NUS
Ph.D. Supervisor: Dr Stéphane Bressan

Team Web & Data Science

MOTIVATION

The main stage for a new generation of cooperative information systems are smart communities such as smart cities and smart nations. In the smart city context in which we position our work, urban planning, development and management authorities and stakeholders need to understand and take into account the mobility patterns of urban dwellers in order to manage the sociological, economic and environmental issues created by the continuing growth of cities and urban population. We address the issue of the detection of communities of commuters which is one of the crucial aspects of smart community analysis.

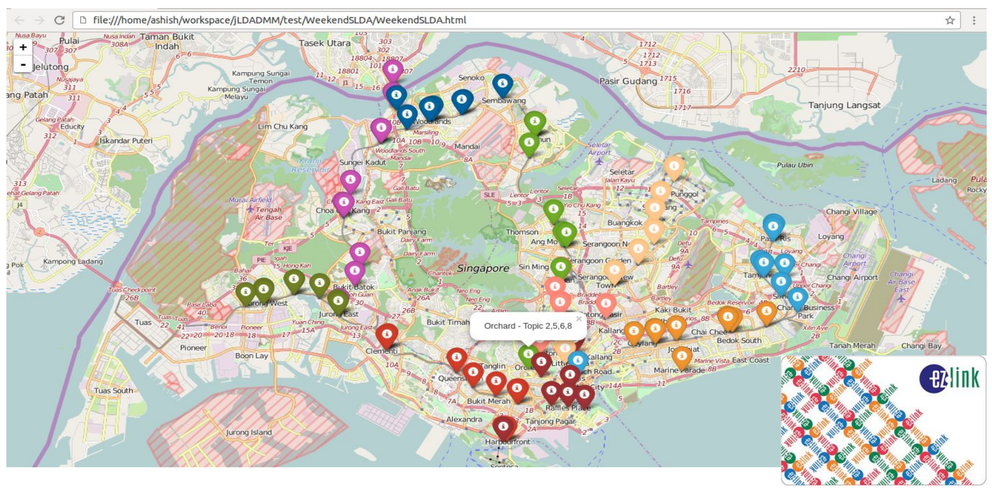

Figure 1: Weekend Spatial Mobility Patterns

GENERATIVE MODELS

- Adapted and Extended Latent Dirichlet Allocation (LDA)[1] to handle Spatio-Temporal Data
- Assumptions:
  - Each document is a probability distribution over *K* topics
  - Each topic is a probability distribution over *N* words

Model	Document	Words
SLDA	Commuter	Places visited by the commuter
TLDA	Commuter	Timestamps at which the commuter visits different places
STLDA	Commuter	Spatio-temporal events of the commuter

Table 1: Analogy between LDA and its spatio-temporal adaption

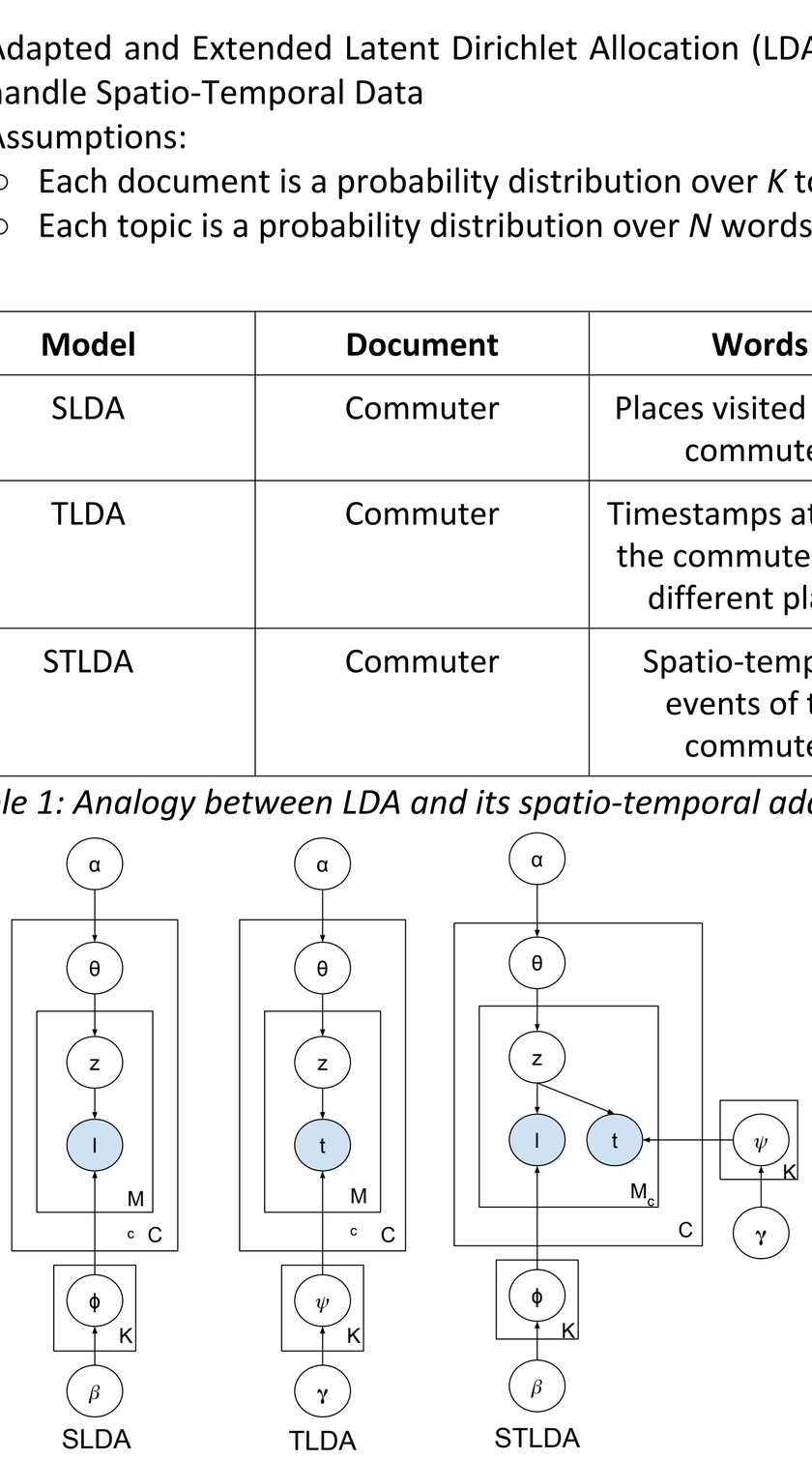

Figure 2: Plate diagrams for the spatio-temporal models

EXPERIMENTS

- Setup:
  - Intel i7 Processor machine with 8GB RAM and inference of the models is written in Java
- Dataset:
  - Real-world public transportation network data
  - Filtered two weekdays and two weekends

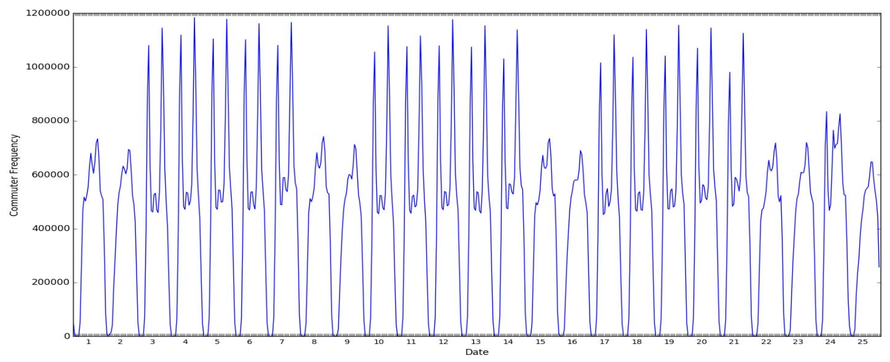

Figure 3: Raw statistics of the data

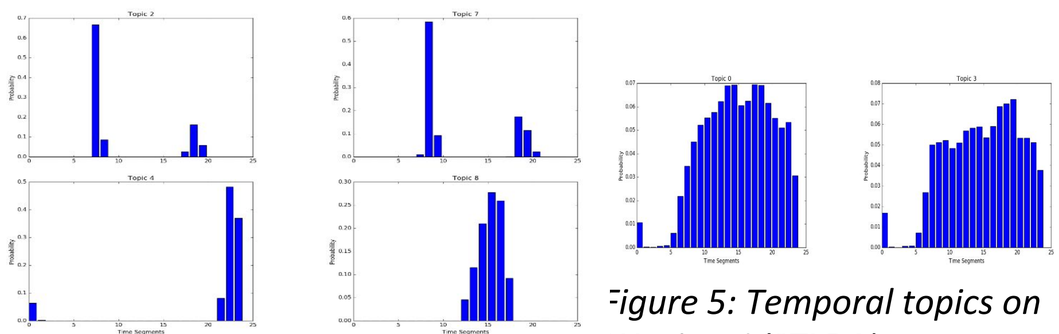

Figure 4: Temporal topics on Weekday (TLDA)

Figure 5: Temporal topics on Weekend (STLDA)

- Proposed method is shown to be more effective and efficient than graph based community detection techniques.

FUTURE WORK

- We are extending the current approach to handle visits which can not be characterized by a known set of fixed location.
- We are looking for Generative Models which can be learn temporal variations in the spatial data to simulate trajectories of objects.

COMMENTS

- This work is a collaboration between NUS, Télécom ParisTech and A*STAR
- This work has been published in CoopIS 2016
- We are aiming to publish the ongoing work in DEXA 2017

REFERENCES

1. Blei, David M., Andrew Y. Ng, and Michael I. Jordan. "Latent dirichlet allocation." *Journal of machine Learning research* 3.Jan (2003): 993-1022.