Comparison of Numerical Integration Methods for Bayesian Inference

Viraj Parimi[2], Aidan Lakshman[3], Isaac K. Isukapati[1], Stephen F. Smith[1]

[1]The Robotics Institute, Carnegie Mellon University

[2]Indraprastha Institute of Information Technology

[3]University of Central Florida

Overview

- Incorporating knowledge of buses into adaptive signal control improves the accuracy schedules
- Real-time bus dwell time prediction is a crucial step in this regard
- Bayesian Hierarchical Model is shown to be effective in constructing a highly predictive model as it requires minimal historical data, is not prone to overfitting, and delivers a confidence level in its predictions.
- Such problems require exact posterior inference which is computationally intractable, however approximate statistical methods such as Markov-Chain Monte Carlo (MCMC) offer schemes for drawing a series of correlated samples that will converge to the target distribution.

Problem

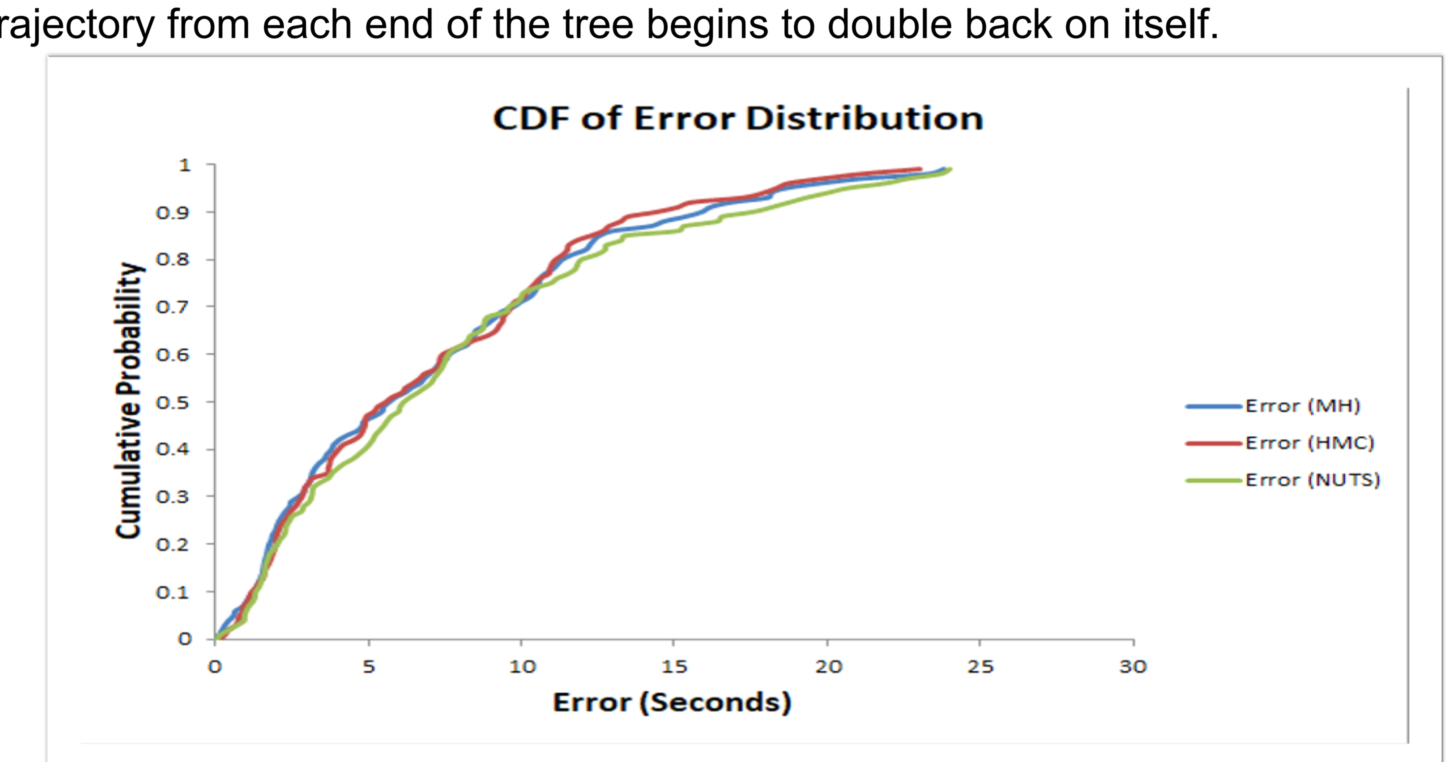
- Numerically approximating the posterior distribution in a computationally efficient manner plays an instrumental role in field implementation of the system.
- In this research, we explore and benchmark three types of numerical integration techniques both on the basis of accuracy and computational efficiency.

Approach

- We analyze the dataset of timestamped records for dwell time and number of onboarding and alighting passengers for buses in the Pittsburgh area.
- Our dataset was drawn from data collected in 2012 from multiple intersections throughout the month of October
- Dwell time distribution was modelled using a Log-Normal distribution with onboarding and alighting information as the covariates. MCMC algorithms were used in conjunction with posterior predictive sampling in order to make predictions on the data. Three MCMC Numerical Integration techniques were tested: **Metropolis Hastings**, **Hamiltonian MC** and **NUTS**

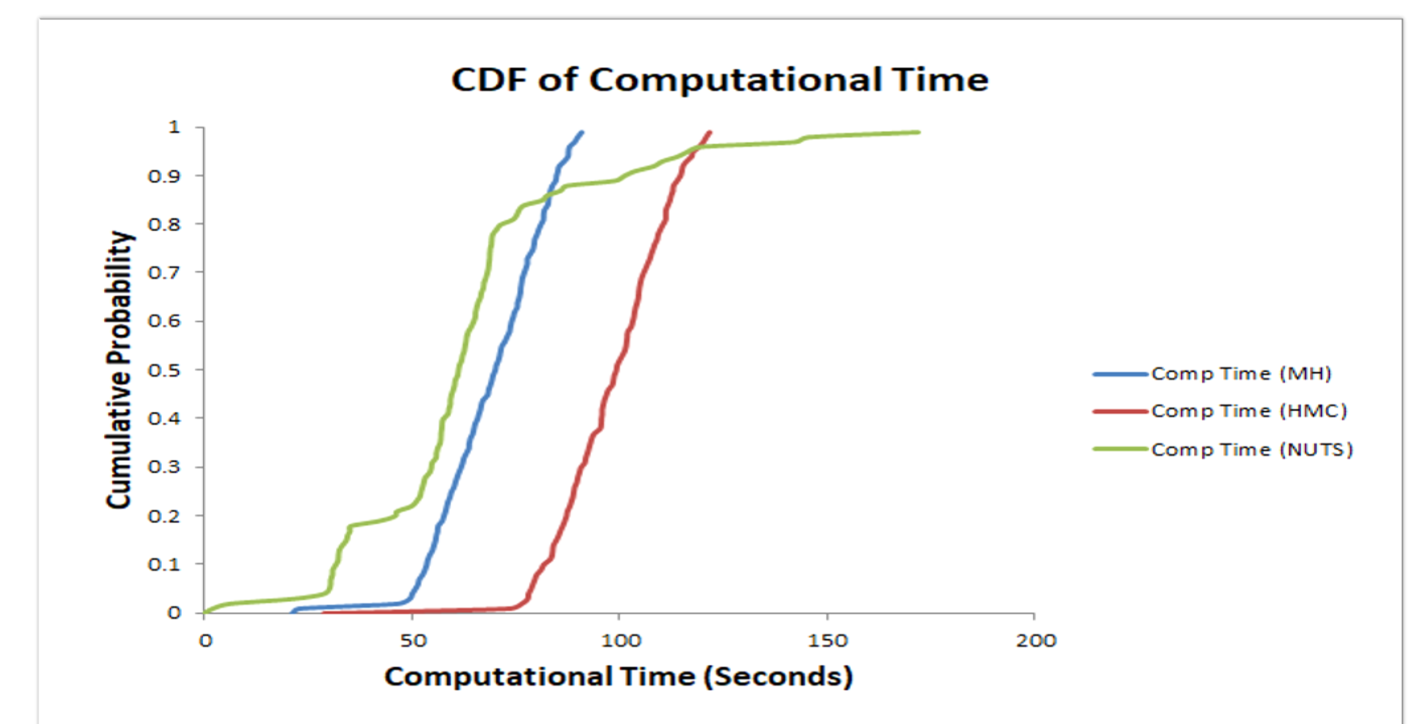
Numerical Integration Techniques

- **Metropolis-Hastings MCMC:** Candidate solutions are randomly chosen from a probability distribution, which are then accepted or rejected as posterior samples based on the full joint density of prior distributions.
- **Hamiltonian MC:** Realizes the solution set as a physical system, then uses a vector field to simulate momentum to move the Markov Chain across the distribution. Can sample much more efficiently than MH, but requires precise parameter tuning.
- **NUTS:** NUTS is a variation of HMC recently developed to fix the difficulties of hand-tuning parameters by looking at the trajectory of the Markov Chain in the variable space. This sampling process generates a balanced binary tree whose leaf nodes consist of position-momentum pairs. The sampling is stopped when the sub-trajectory from each end of the tree begins to double back on itself.



Cumulative Density Function of Error Distribution for Center and Aiken bus-stop

Type of Sampler	Percentile Error				% of predictions in [-5, 5]
	15th	50th	75th	95th	
Metropolis Hastings	1.56	5.68	10.57	18.64	46.5
Hamiltonian MC	1.66	5.58	10.46	18.35	47.3
NUTS	1.6	6.13	10.95	29.68	41.3



Cumulative Density Function of Computational Time for Center and Aiken bus-stop

Type of Sampler	Percentile Computation Time			
	15th	50th	75th	95th
Metropolis Hastings	55.57	69.91	79.04	87.53
Hamiltonian MC	84.77	99.33	108.25	117.89
NUTS	34.3	61.33	68.99	116.57

Conclusions

- Based on the results we see that with regards to the error distribution the HMC algorithm works marginally better compared to others but at the cost of computation inefficiency.
- We also observed that NUTS was computationally efficient within the 75th percentile but at the cost of higher inaccuracy compared to other algorithms
- This shows the tradeoff between computational inefficiency and accuracy of the model

Future Work

- We plan to extend the comparison and work with PyMC3 to get more detailed insights
- We plan to implement this optimized model into Surtrac adaptive signal controller algorithm. This data will be used by the model to increase traffic flow efficiency in intersections involving transit vehicles
- This enhanced algorithm will be implemented and tested throughout Pittsburgh

Acknowledgements

Thank you to Dr. Isaac Isukapati, Dr. Stephen Smith for their constant support and assistance with this project. Also, thank you to Rachel Burcin and Dr. John Dolan for organizing this program and for setting up workshops and events for us throughout the summer. Finally, thank you to NSF and FICCI, who made it possible for us to participate this summer.