



Pós-graduação em Computação Aplicada
Universidade Estadual de Feira de Santana (UEFS)

S2I+: Armazenamento Eficiente em Índice Espaço-Textual

Tiago F. Athayde-Novaes, Felipe L. Fonseca e João B. Rocha-Junior

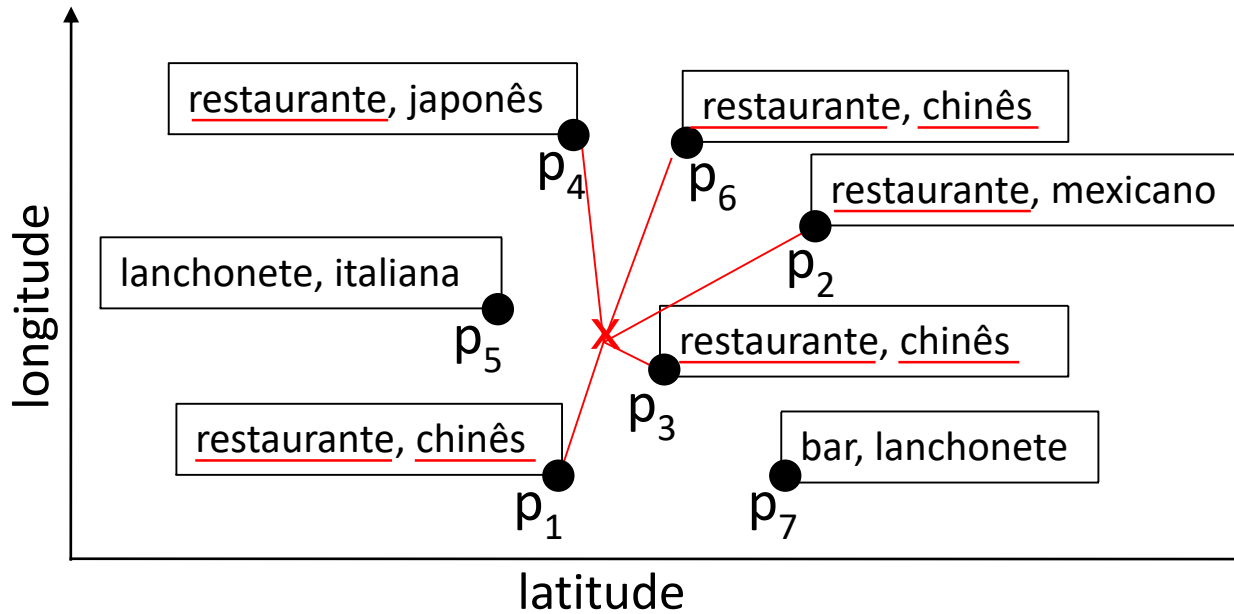
Outline

- Introduction;
- Top- k Spatio-Textual Query;
- Spatial Inverted Index (S2I);
- Motivation;
- Data Storage Study;
- Experimental Evaluation;
- Conclusion.

Introduction

- Popularization of spatio-textual objects;
 - Example: objects of twitter sent from a smartphone with a GPS.
- Top- k Spatio-Textual Query;
 - Parameters: spatial location, query keywords and k ;
 - Result: the k best spatial-textual objects.

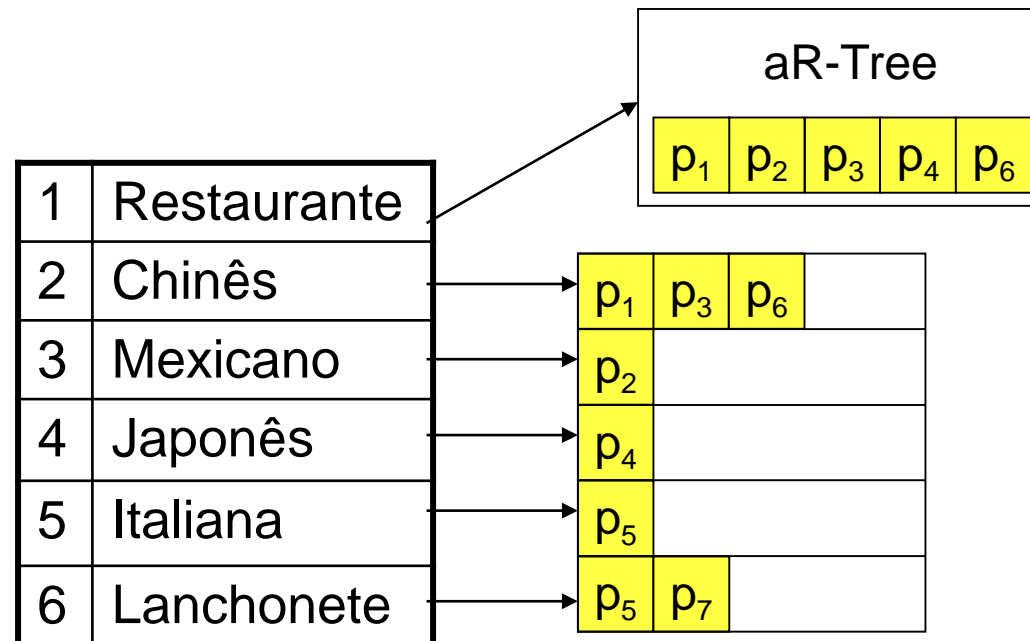
Top- k Spatio-Textual Query



- Spatial location = x , query keywords = **restaurante chinês** and $k = 2$;
- Result p_3 e p_1 .

Spatial Inverted Index - S2I

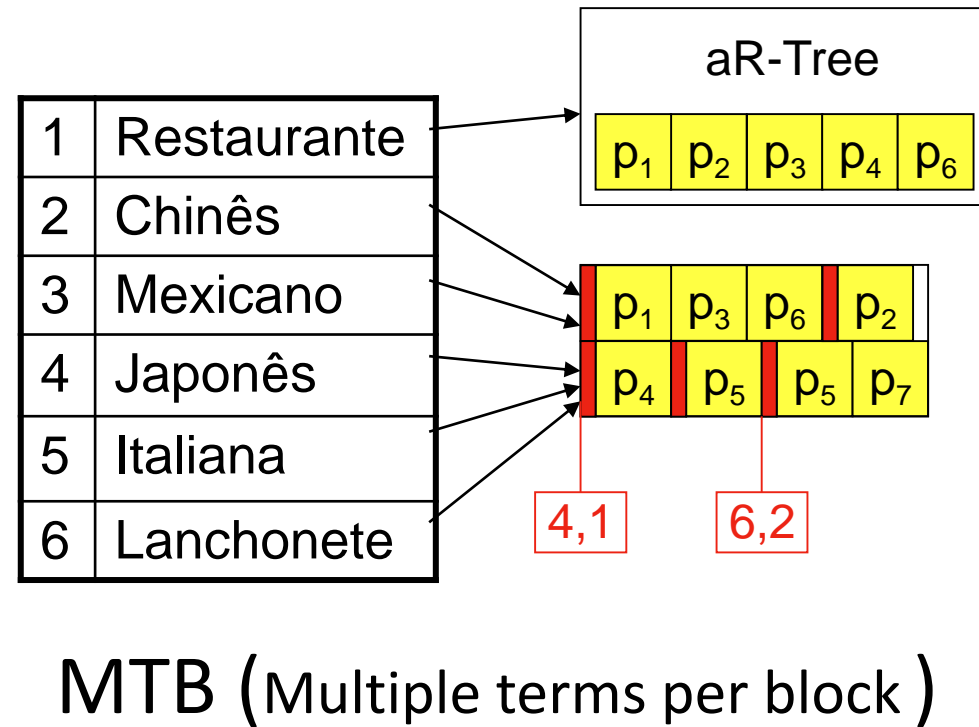
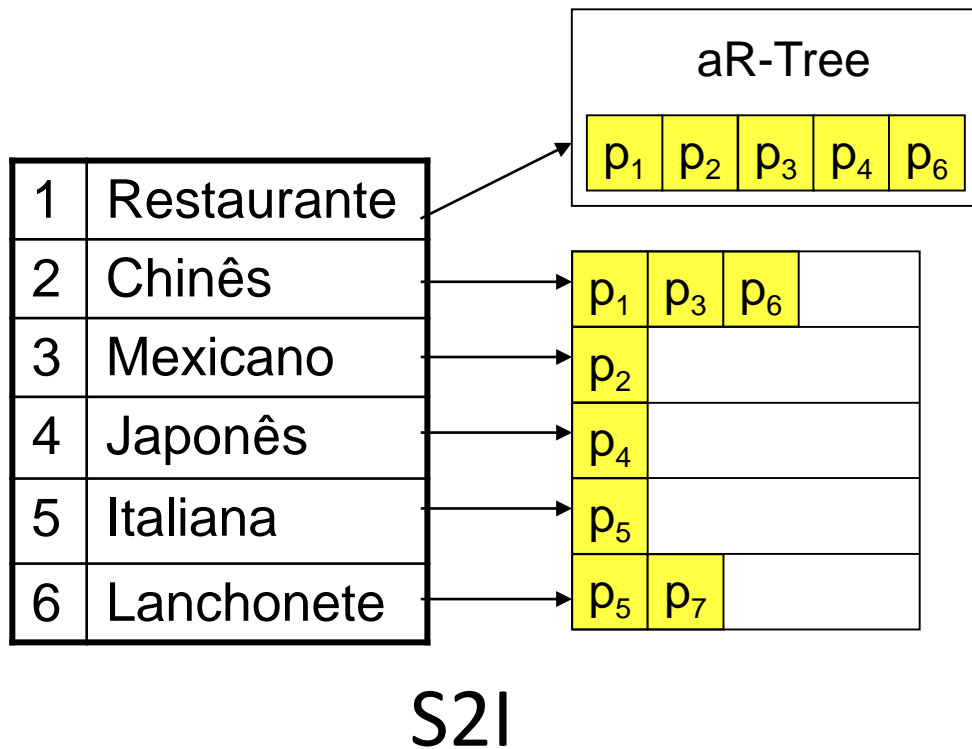
- Similarly to an inverted index;
 - Blocks for infrequent terms;
 - Index Multidimensional (aR-Tree) for frequent terms.



Motivation

- Needs a large amount of disk space (Chen et al 2013);
- Reduces the space consumed by S2I;
- Study data storage aspects in S2I;
- Propose improvements in terms of data storage organization.

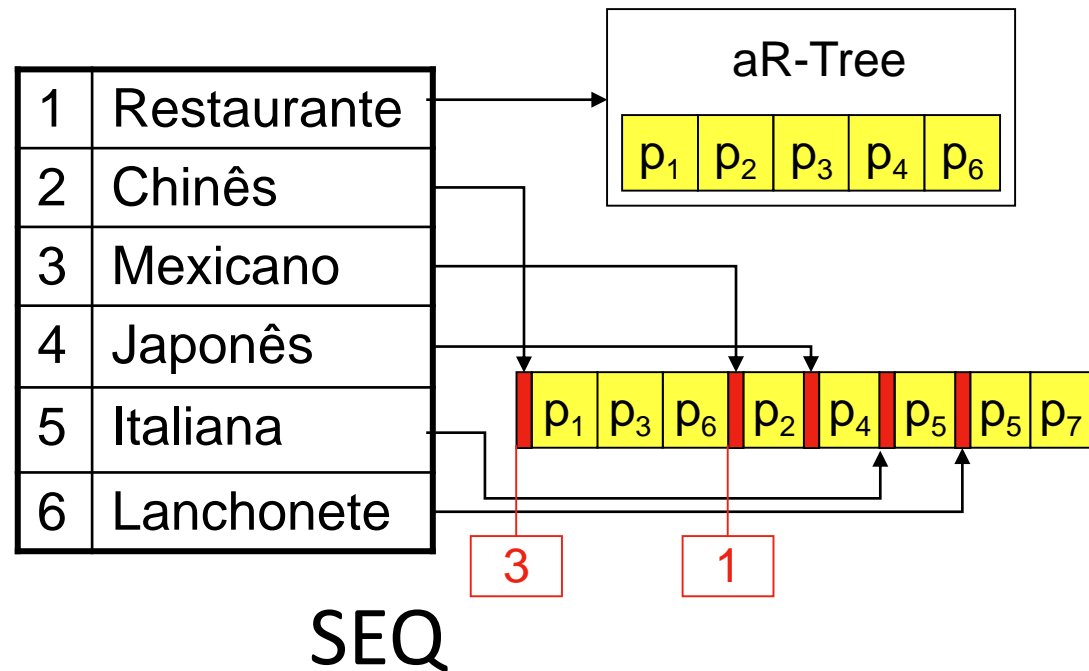
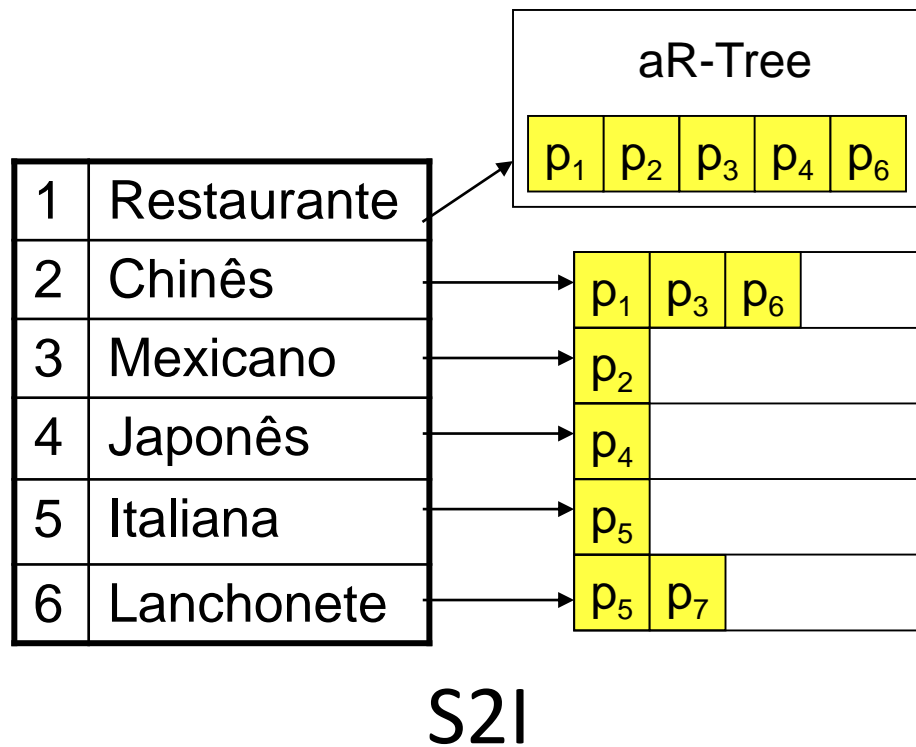
Data Storage Study



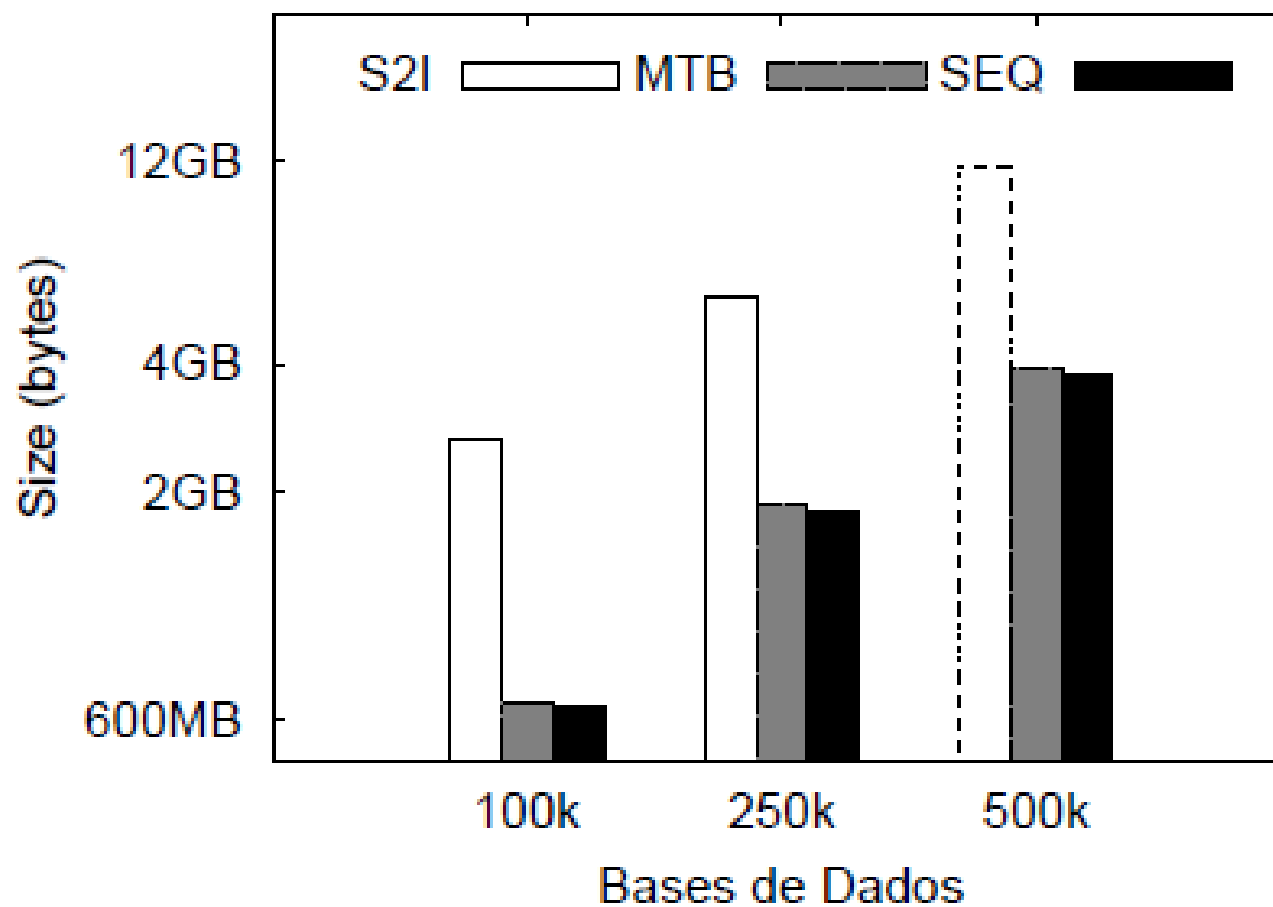
- Allows objects of infrequent terms occupy the same block.

Sequential Storage (SEQ)

- The objects are stored in a single file with "blocks" of varying sizes.

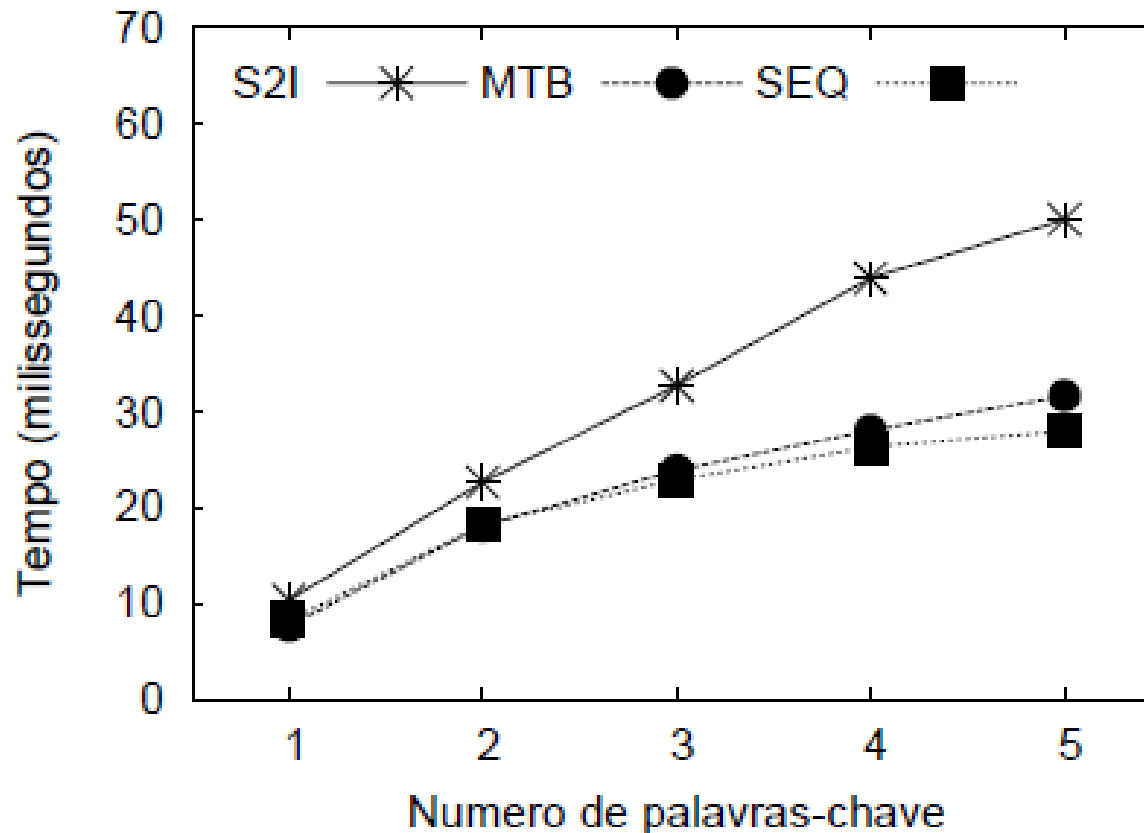


Experiment (build)



Experiment (queries)

- 800 queries with random location and query keywords. The $k = 10$.



Conclusion

- The proposed approaches MTB and SEQ reduce the size of S2I;
- The proposed approaches also present better performance in terms of response time.

Thanks!

More information...

Tiago F. Athayde-Novae
tiagoathayde@hotmail.com