



# Location-Based Instant Search

Shengyue Ji\* Chen Li

Google, Inc University of California, Irvine

Work done in University of California, Irvine.

# Location-Based Keyword Search

- Examples

- Finding movie theaters in Manhattan on Google Maps
- Finding Japanese restaurants near Disneyland in

The image shows a screenshot of a Google Maps search for "movie theaters in manhattan" with a Yelp overlay. The Yelp overlay is for a search of "japanese restaurant" near "disneyland in anaheim". The search results are filtered to show 1 to 10 of 416 results. The results are sorted by "Best Match" and include filters for "Cities" (Disneyland, Tustin, Orange, Garden Grove), "Distance" (Bird's-eye View, Driving, Biking, Walking, Within 4 blocks), "Features" (Open Now, Good for Groups, Good for Kids, Take-out), "Price" (\$\$\$, \$\$, \$), and "Category" (Japanese, Sushi Bars, Korean, Fast Food).

Rank	Name	Address	Phone	Reviews
1	Onshimo Japanese Cuisine	1956 N Tustin St Orange, CA 92665	(714) 898-0098	773 reviews
2	Orange Sushi	311 N Euclid St Fullerton, CA 92832	(714) 525-1387	28 reviews
3	Yuki Sushi	13601 Beach Blvd Stanton, CA 90680	(714) 903-8899	199 reviews
4	Sushi Wasabi	14400 Newport Ave Tustin, CA 92780	(714) 505-3498	477 reviews

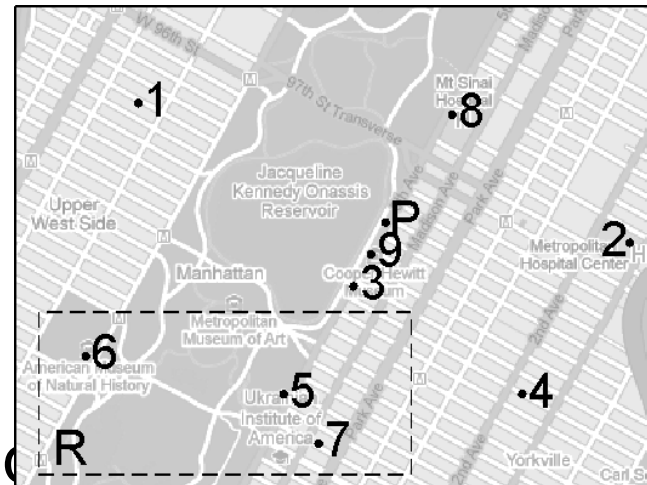
# Outline

- Location-Based Keyword Search
  - Existing Methods
  - Improvement: FEH
- Location-Based Instant Keyword Search
  - Representing Prefixes on FEH
  - Compressing Prefixes

# The Problem

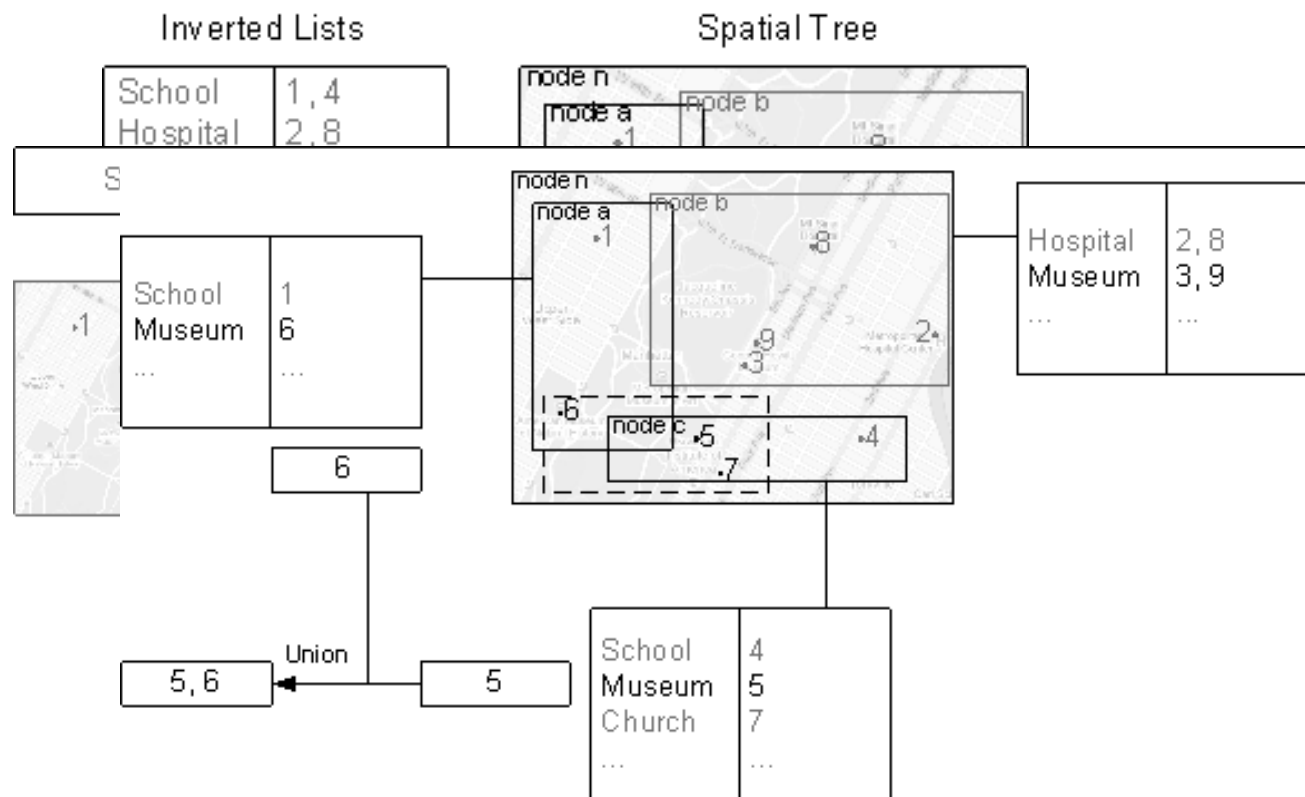
- Data

- | ID  | Name                               | Location            |
|-----|------------------------------------|---------------------|
| 1   | Trinity School                     | <40.7903, -73.9707> |
| 2   | Metropolitan Hospital Center       | <40.7846, -73.9441> |
| 3   | Solomon R. Guggenheim Museum       | <40.7831, -73.9596> |
| 4   | Brearly School                     | <40.7783, -73.9501> |
| 5   | Metropolitan Museum of Art         | <40.7786, -73.9629> |
| 6   | American Museum of Natural History | <40.7791, -73.9730> |
| 7   | Manhattan Church of Christ         | <40.7766, -73.9613> |
| 8   | Mt Sinai Hospital                  | <40.7901, -73.9538> |
| 9   | Cooper Hewitt Museum               | <40.7844, -73.9580> |
| ... | ...                                | ...                 |



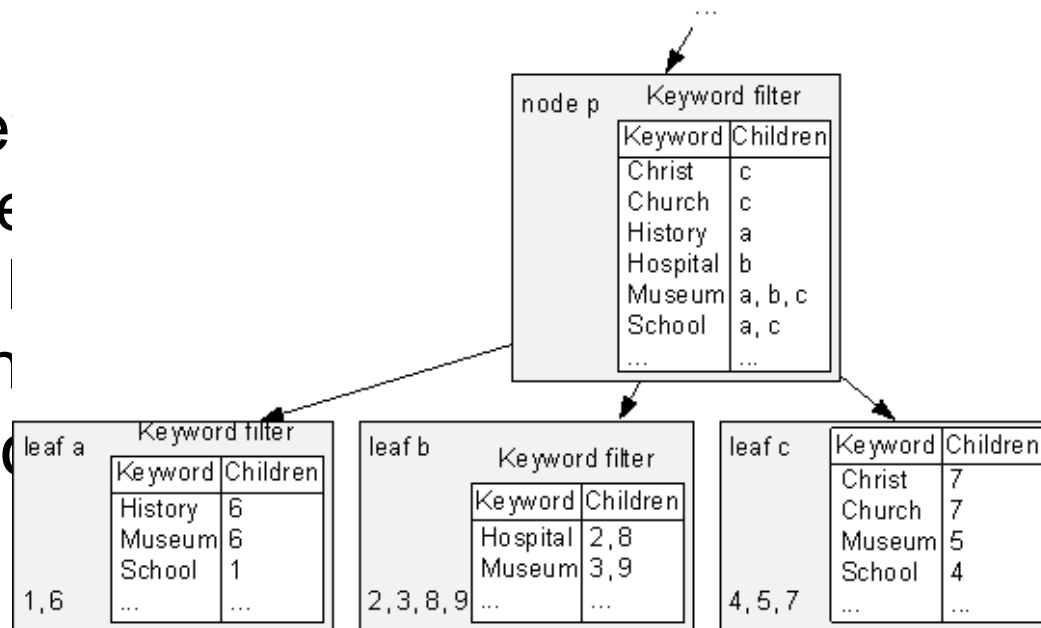
# Existing Methods

- Using separate indices
  - Inverted file and R\*-tree double index
  - First inverted file then R\*-tree
  - First R\*-tree then inverted file



# Existing Methods

- Augmenting R\*-tree with keyword filters
- Variations
  - Set of keywords
  - Signature
  - Inverted index
  - Min-hash
- Inside vs. outside



# FEH: Filtering-Effective Hybrid Index

- Intuition
  - Index creation time: treating the keywords at a node differently.
  - Query time: directly accessing the records for very selective keywords.

Selectiveness

Record selectiveness

No more than  $tr$  records in  $n$  have  $w$ .

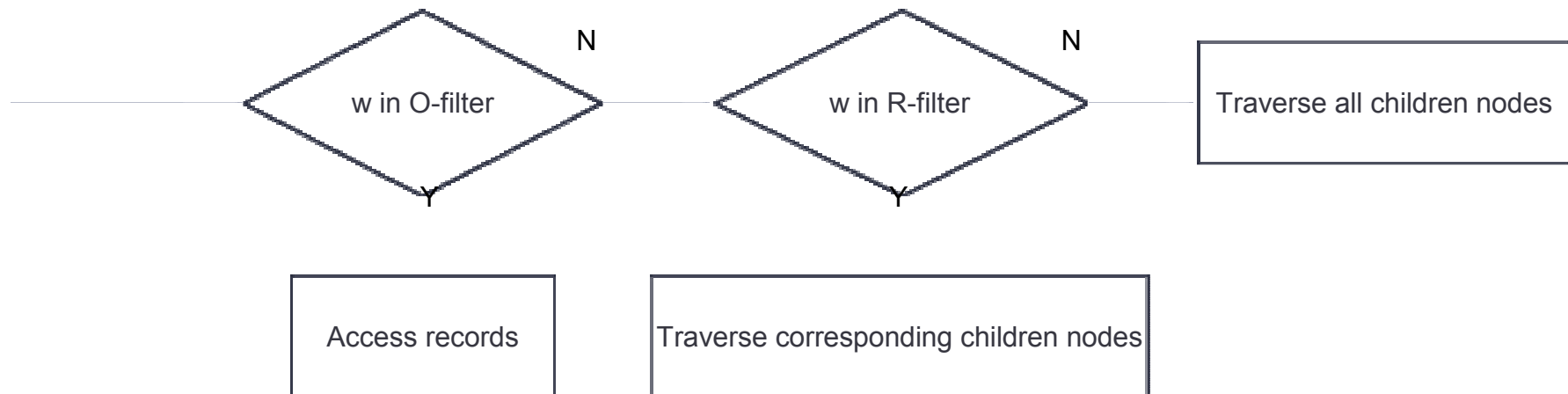
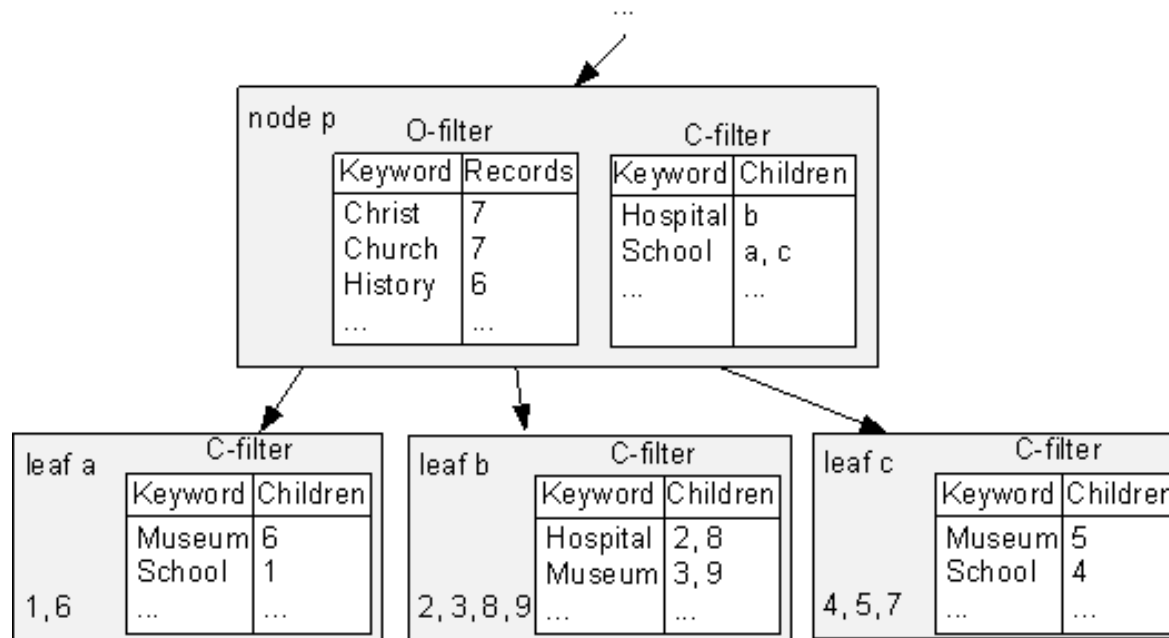
Child selectiveness

Node  $n$  has at least one child that does not have  $w$ .

Non-selectiveness

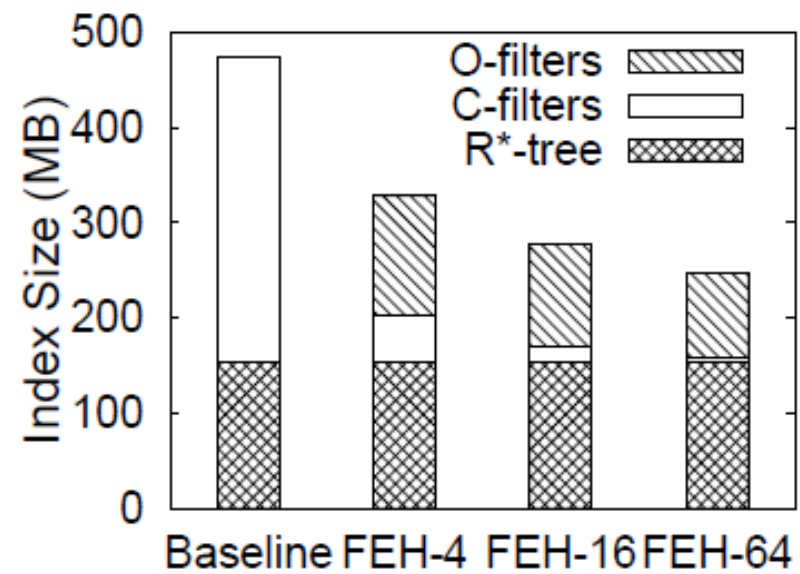
Otherwise.

# FEH: Example

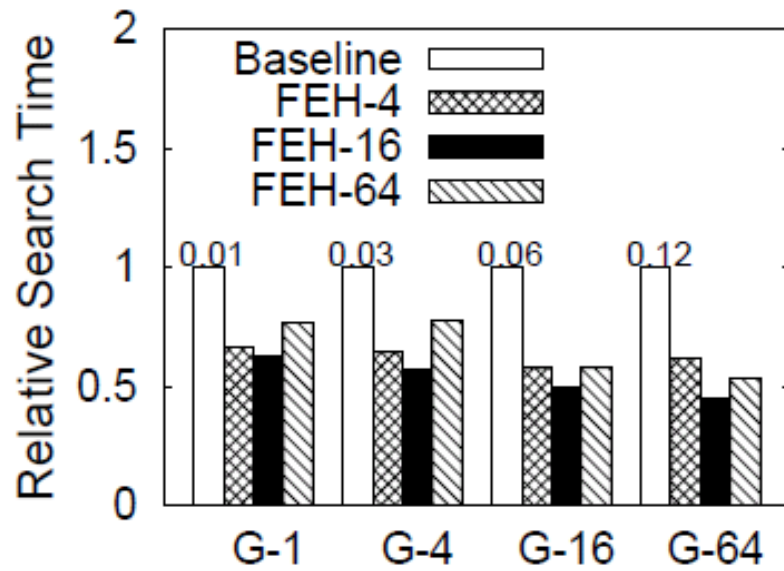




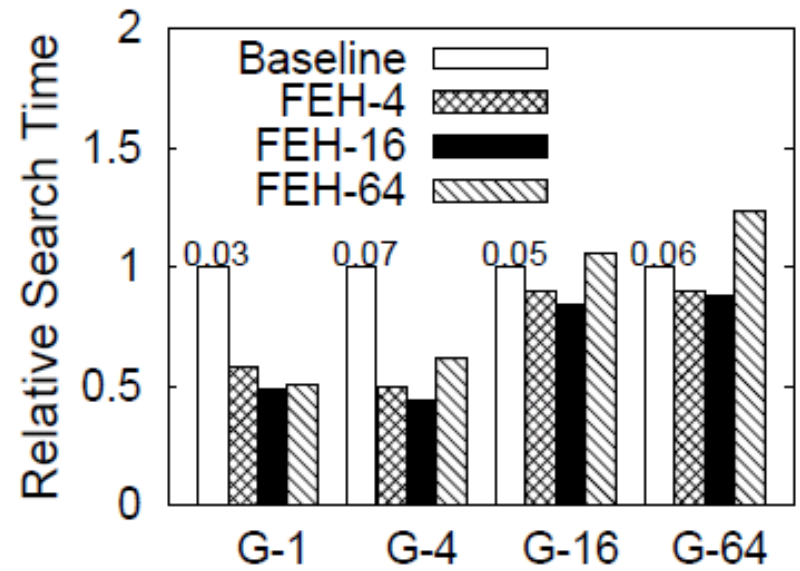
# FEH: Experiments



# FEH: Experiments



(a) Range queries



(b) kNN queries ( $k = 10$ )

# Outline

- Location-Based Keyword Search
  - Existing Methods
  - Improvement: FEH
- **Location-Based Instant Keyword Search**
  - **Representing Prefixes on FEH**
  - **Compressing Prefixes**

# Location-based Instant Keyword Search

- Query
  - Range query  $\langle R, W \rangle$
  - kNN query  $\langle P, W \rangle$
  - $W = \{w_1, w_2, \dots, w_l\}$

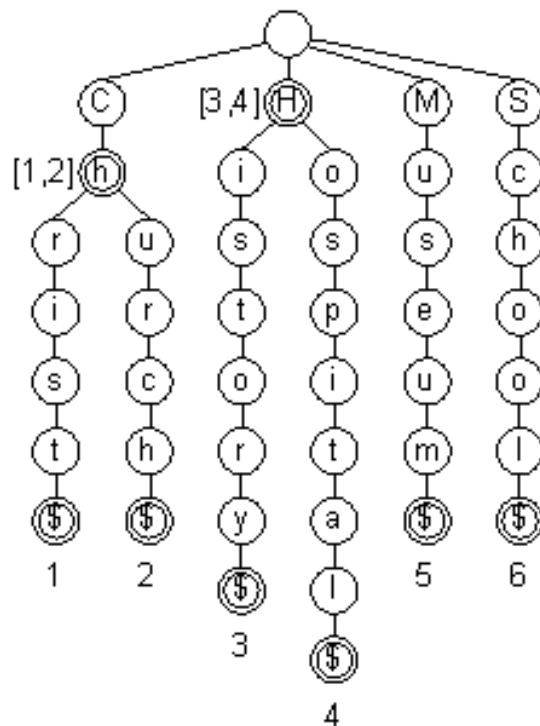
The screenshot displays a search interface with three main sections: a search input field, a list of results, and a map. The search input field contains 'Metropolitan Mus' and 'Manhattan'. The list of results shows various entries related to the Metropolitan Museum of Art and Music Co, each with a red circular icon and a count in parentheses. The map shows the New York City area with several red circular markers labeled A through I. A callout box for the Metropolitan Museum of Art is open, showing its address, phone number, and coordinates.

Business	Location	Map Search Demo
Metropolitan Mus	Manhattan	
<b>A</b> Metropolitan Museum of Art (3)		<b>Metropolitan Museum Of Art</b> 5th Ave & 82nd New York, NY 10021 (212) 879-5500 [40.7671, -73.959]
<b>B</b> Metropolitan Museum Art Shop (2)		
<b>C</b> Metropolitan Museum Cf Art (8)		
<b>D</b> Metropolitan Museum Art Gift		
<b>E</b> Metropolitan Museum-Art		
<b>F</b> Metropolitan Museum-Art Store		
<b>G</b> Metropolitan Museum-Art Str		
<b>H</b> Metropolitan Music		
<b>I</b> Metropolitan Music Co (2)		

# Representing Prefix Filters

- Filter explosion

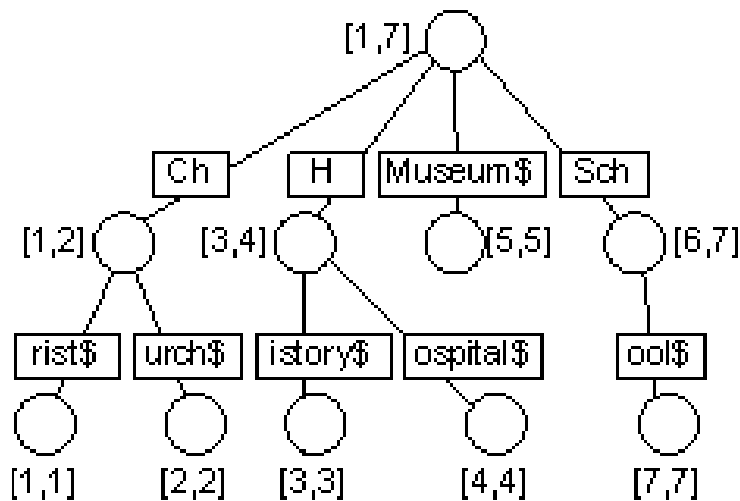
node p		Keyword filter	
		Keyword	Children
		Christ	c
		Church	c
		History	a
		Hospital	b
		Museum	a, b, c
		School	a, c
		...	...



Prefix Filter			Ho	[4,4]	b
C	[1,2]	c	Hos	[4,4]	b
Ch	[1,2]	c	Hosp	[4,4]	b
Chr	[1,1]	c	Hospi	[4,4]	b
Chri	[1,1]	c	Hospit	[4,4]	b
Chris	[1,1]	c	Hospita	[4,4]	b
Christ	[1,1]	c	Hospital	[4,4]	b
Christ\$	[1,1]	c	<b>Hospital\$</b>	<b>[4,4]</b>	<b>b</b>
Chu	[2,2]	c	M	[5,5]	a,b,c
Chur	[2,2]	c	Mu	[5,5]	a,b,c
Churc	[2,2]	c	Mus	[5,5]	a,b,c
Church	[2,2]	c	Muse	[5,5]	a,b,c
Church\$	[2,2]	c	Museu	[5,5]	a,b,c
H	[3,4]	b,c	Museum	[5,5]	a,b,c
Hi	[3,3]	c	<b>Museum\$</b>	<b>[5,5]</b>	<b>a,b,c</b>
His	[3,3]	c	S	[6,6]	a,c
Hist	[3,3]	c	Sc	[6,6]	a,c
Histo	[3,3]	c	Sch	[6,6]	a,c
Histor	[3,3]	c	Scho	[6,6]	a,c
History	[3,3]	c	Schoo	[6,6]	a,c
History\$	[3,3]	c	School	[6,6]	a,c
			<b>School\$</b>	<b>[6,6]</b>	<b>a,c</b>

# Representing Prefix Filters

- Using a radix trie
- Store prefixes as intervals



Logical representation of consistent prefixes

Prefix Filter

	Interval	Children of R*-tree node p
	[1,7]	a, b, c
Ch	[1,2]	c
Christ\$	[1,1]	c
Church\$	[2,2]	c
H	[3,4]	b, c
History\$	[3,3]	c
Hospital\$	[4,4]	b
Museum\$	[5,5]	a, b, c
Sch	[6,7]	a, c
School\$	[7,7]	a, c

Physical representation

# Representing Prefix Filters

- Why intervals?
  - A radix trie node  $\leftrightarrow$  an interval
  - Easy to test whether one interval is an ancestor of another
  - Can represent the lexicographical order

$I_1 = [Id_{min1}, Id_{max1}]$  and  $I_2 = [Id_{min2}, Id_{max2}]$ , we define:

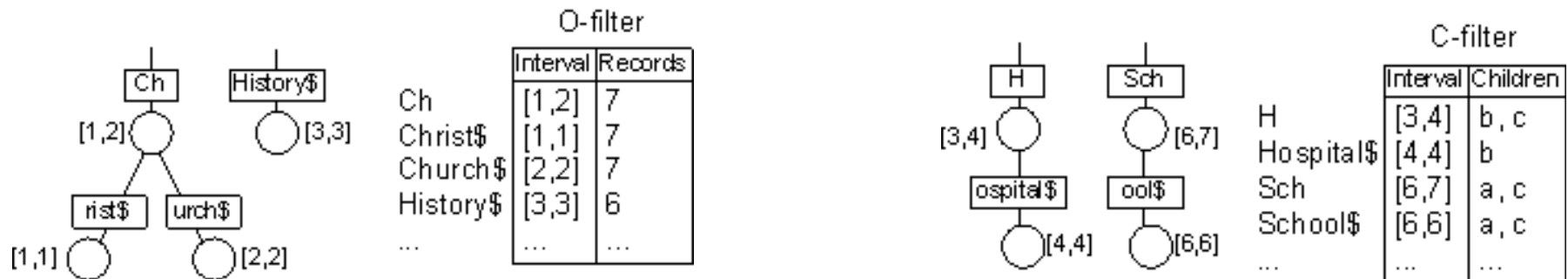
$$\left\{ \begin{array}{l} I_1 < I_2, \text{ if } Id_{min1} < Id_{min2} \vee (Id_{min1} = Id_{min2} \\ \quad \wedge Id_{max1} > Id_{max2}); \\ I_1 = I_2, \text{ if } Id_{min1} = Id_{min2} \wedge Id_{max1} = Id_{max2}; \\ I_1 > I_2, \text{ otherwise.} \end{array} \right.$$

# Representing Prefix Filters

- On FEH

node p

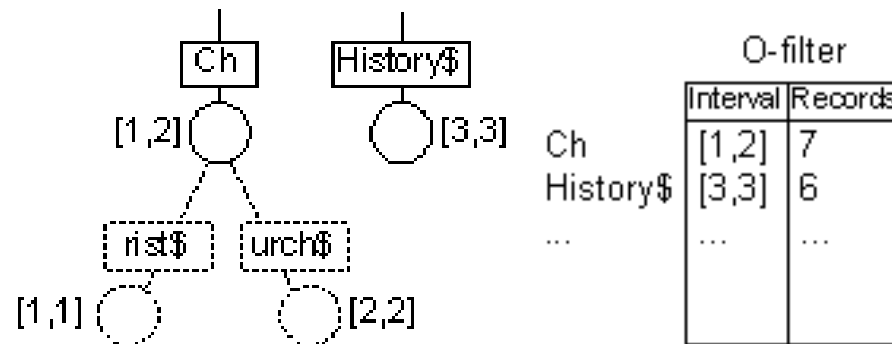
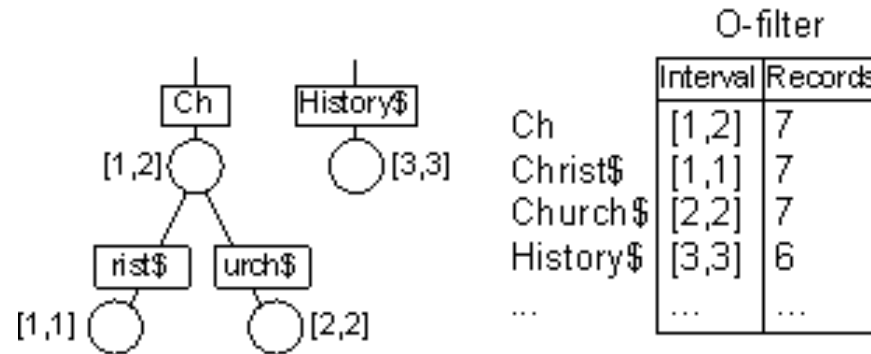
O-filter		C-filter	
Keyword	Records	Keyword	Children
Christ	7	Hospital	b
Church	7	School	a, c
History	6	...	...
...	...	...	...



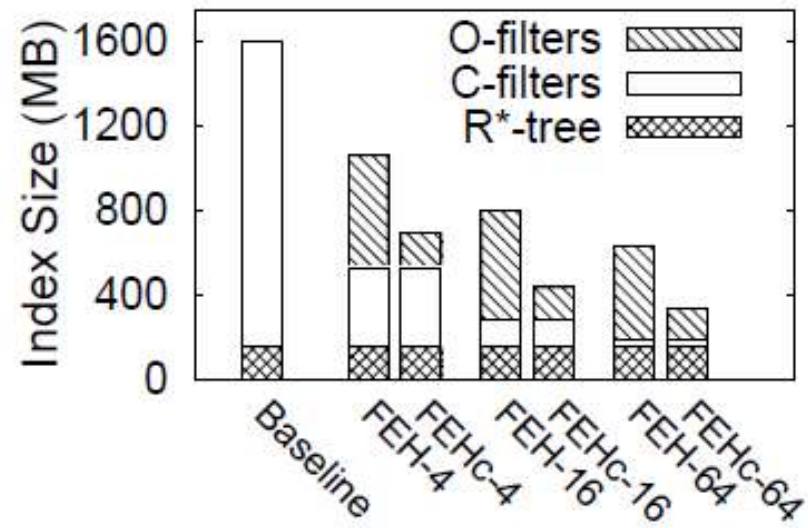


# Compressing Prefix Filters

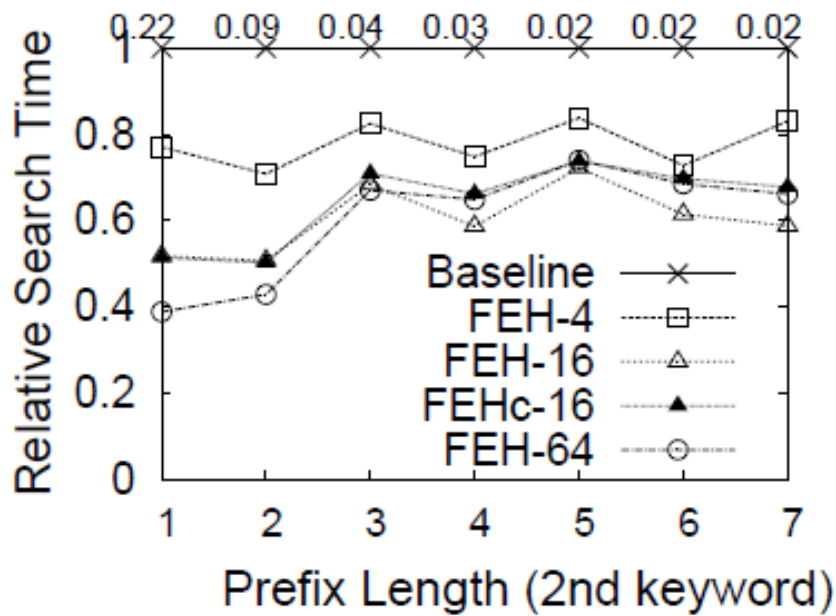
- Compress O-Filters
  - Using ancestors to represent descendents
  - Searching for ancestors in the filter



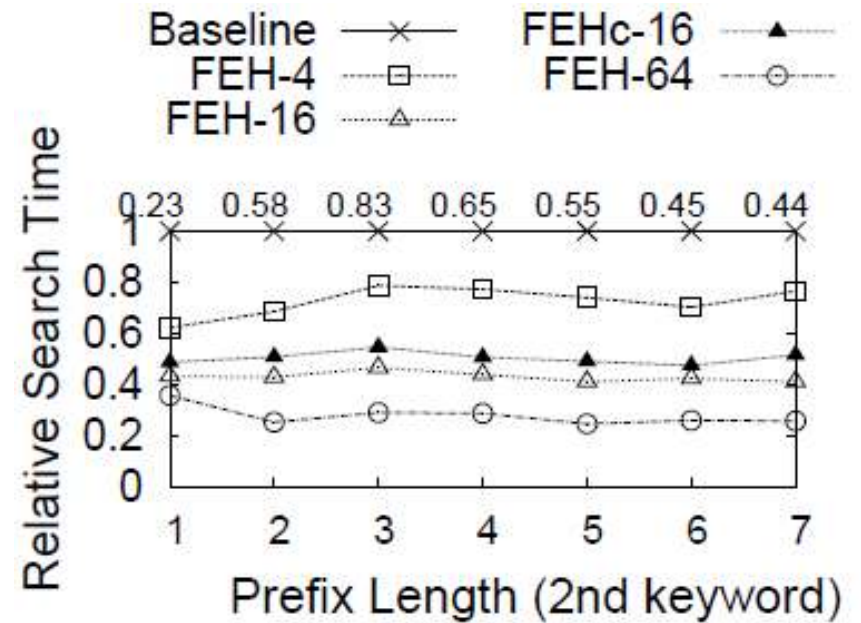
# Experiments



# Experiments



(a) Range queries



(b) kNN queries

THANK YOU!

Questions?