

# GeoHash Retrieval

Thursday, March 21, 2024 8:45 AM

## GeoHash Retrieval:

- ① user may be interested in items in their neighborhood.
- ② **GeoHash**: "embedding" of longitude and latitude.  
correspond to a rectangle area on map
- ③ **index**: GeoHash  $\rightarrow$  popular items (from new to old)
- ④ No **personalization** in this retrieval.  
merely based on GeoHash code.

## Visualization:

GeoHash		The newest $k$ high-quality items
ABCDE	$\longrightarrow$	item #1, ... item #k
GACVC	$\longrightarrow$	item #1, ... item #k
$\vdots$		$\vdots$
CFAGX	$\longrightarrow$	item #1, ... item #k

Similar concept can be used for "city-based retrieval"  
Instead of using "GeoHash" code, use "city" instead.

# Author-Based Retrieval

Thursday, March 21, 2024 8:45 AM

## Author-Based Retrieval:

- ① users follow some authors (influencer)
- ② Index 1: user  $\rightarrow$  following authors  
author  $\rightarrow$  published posts  
("author" is also "user")
- ③ Retrieval: user  $\rightarrow$  following authors  $\rightarrow$  newly published posts

## Interaction Author Retrieval:

- ① user likes a post, then potentially like other notes published by the same author
- ② index: user  $\rightarrow$  interacted authors
- ③ Retrieval: user  $\rightarrow$  interacted authors  $\rightarrow$  most recent posts

## Similar Author Retrieval:

- ① if user likes an author, then they must like similar authors
- ② index: author  $\rightarrow$  k similar authors
- ③ retrieval: user  $\rightarrow$  potentially like authors  $\rightarrow$  similar authors  
(n authors) (n.k authors)  
most recent notes  $\leftarrow$   
(n.k notes)



## Cache - Based Retrieval:

- ① some items were retrieved and ranked, but NOT exposed to users because of "re-ordering"
- ② put top 50 items but not exposed into cache and used as retrieval results
- ③ cache has a fixed size: need "pop out" mechanism
  - (a) once exposed to user, remove it
  - (b) if cache limit is reached, remove the oldest item.
  - (c) each item can only be retrieved  $N$  times after  $N$  times. remove it
  - (d) each item is stored  $X$  days. after  $X$  days. remove it.