Constrained ANN

e.g., Select * from A
where 0600 ≤ time ≤ 1000
ORDER By embedding <-> [-2,5...]
Limit 10

Constraints are defined on Traditional attrib.
- Categorical (Label, ...) Non-ordinal
- Numeric (Time, Location, ...) Ordinal

- Single-Predicate: the Constraint is on one attribute (e.g., Time ≥ 1000)
- Multi-Predicate: " " Can be on multiple attributes:

Open Problem

where
time ≥ 1000
AND
Price ≤ 5000
AND
Label = 'x'

Existing work: mostly Single Predicate
- Numeric ⇒ Range Predicates → SeRF, iRangeGraph
- Categorical ⇒ Label " : Filtered DiskANN ...
- both ⇒ General Case: Faiss, Milvus ...

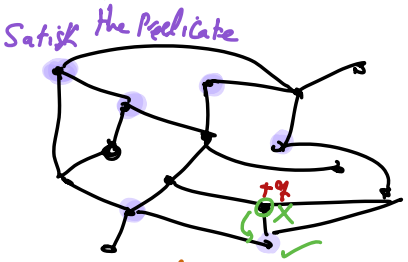
Approaches
- Prefiltering: First Filter, then Search
- Post " : " Search, " Filtering
- In " : Filter while Search

Prefiltering
① Find the Tuples that Satisfy the predicate
② Among the ones that didn't prune, Find the ANN.

Issue:
① Don't build an Index ↖ Slow ✓
② Build one for every Possible Index ⇐ Huge Space ✗

effective when only a Small number of tuples Pass the Predicate

Post filtering:
keep Generating the Next nearest neighbors (e.g., Using HNSW graph) and reject the result if it does not Satisfy the Condition



Not effective if Selectivity is low!

Pre filtering : Good when Selectivity is low

Post " : " " " "
is high.

⇒ Hybrid? | if low → Pre filtering
" high → Post "

what if $\Theta(\sqrt{n})$ tuples pass the Predicate? X

⇓

In Filtering:

Build Indices that enable Filtering during the Search.

SeRF: Single Range Predicates

e.g., $l \le time \le h \Rightarrow time \in [l, h]$

Core Idea:

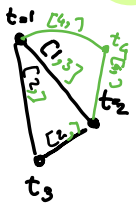
Build One HNSW graph for half-bounded Ranges:

$[1, \dots, X]$

Observation: If we Sort the points based on the filtering attribute (e.g., time), & insert them to the HNSW graph in that order, the set of edges inserted before X, for G_X: The graph for range $[1, X]$

- Segment Graph:

A HNSW graph where edges have a range $[a, b]$, where they were added at iteration a and removed at b



for a half-bound range $[1, t]$, only use the valid edges i.e., $t \in [a, b]$ during the Greedy Search

---

Extending to full ranges:

Build a Segment graph For every Starting Point

G_1 : Segment graph Starting at 1
G_2, " " " 2
⋮
G_K, " " " K

to answer $[a, b]$, use G_a,

⇒ Space Complexity:

n graphs, each $\Theta(n \log n)$

$\Rightarrow \Theta(n^2 \log n)$