

Paper Title

Jaouhara Chanchaf
UM6P

Karima Echihabi
UM6P

Algorithm 2: HEURISTICKNNSEARCH

Input: A query vector q , k , N_{root} .
Output: k Nearest Vectors to q .

```

1 Queue  $Knn[k] \leftarrow \{(null, \infty_1), \dots, (null, \infty_k)\}$ ;
2 Node  $N \leftarrow N_{root}$ ;
3 while  $\neg N.IsLeaf()$  do
4    $SP = N.SplitPolicy()$ ;
5    $N = N.RouteToChildNode(q, SP)$ ;
6  $(v', bsf) = Knn[k]$ ;
7 foreach  $v \in N.Vectors()$  do
8   if  $D(v, q) \leq bsf$  then
9      $Knn.SortedInsert(v, D(v, q))$ ;
10   $(v', bsf) = Knn[k]$ ;
11 return  $Knn$ ;
```

Algorithm 3: KASHIF: INITWORKERTHREAD

Input: $JobArray$, k , $JobCounter$, $Barrier$, $KnnResults$, $AllWorkersDone$.

```

1 while True do
2    $JobIdx = SyncFetchAndAdd(JobCounter)$ ;
3   if  $JobIdx > JobArray.Size()$  then
4     while  $\neg AtomicLoad(AllWorkersDone)$  do
5       Thread reaches Barrier;
6     break; // no more jobs to do, and all workers
           finished their work
7    $q = JobArray[JobIdx]$ ;
8   EXACTKNNSEARCH( $q, k, KnnResults[JobIdx], Barrier$ );
9   if  $JobIdx == JobArray.Size()$  then
10    // completed last job
    AtomicStore( $AllWorkersDone, True$ );
```

Algorithm 4: EXACTKNNSEARCH

Input: A query vectors q , k , $knnResults$, $Barrier$, N_{root} .

```

// local knn results, ordered by distance
1 Queue  $Knn[k] \leftarrow \{(null, +\infty_1), \dots, (null, +\infty_k)\}$ ;
2 Queue  $pq \leftarrow \{\}$ ; // priority queue
/* last NN returned by incremental search is at
   position LastIncrResult - 1 */
3 Integer  $LastIncrResult \leftarrow 0$ ;
4 Integer  $UpdateStart \leftarrow 0$ ; // where new results start
5 Integer  $UpdateEnd \leftarrow 0$ ; // where new results end
6 Boolean  $NewIncrement \leftarrow False$ ;

/* perform heuristic search and update knn results
   in global array */
7  $KnnResults \leftarrow HEURISTICKNNSEARCH(q, k)$ ;
// initialize priority queues
8  $pq.Add(N_{root})$ ;

9 while  $\neg pq.Empty()$  do
10   $NewIncrement = False$ ;
11  while  $\neg NewIncrement$  and  $\neg pq.Empty()$  do
12     $N \leftarrow pq.Pop()$ ;
13    for  $pos \leftarrow LastIncrResult$  to  $k - 1$  do
14       $(v', bsf) \leftarrow Knn[pos]$ ;
15      if  $D_{lb}(N, q) > bsf$  then
16         $LastIncrResult \leftarrow pos + 1$ ;
17         $NewIncrement \leftarrow True$ ;
18    if  $NewIncrement$  then
19       $UpdateStart \leftarrow UpdateEnd$ ;
20       $UpdateEnd \leftarrow LastIncrResult$ ;
21    if  $N.IsLeaf()$  then
22       $(v', bsf) \leftarrow Knn[k]$ ;
23      foreach  $v \in N.Vectors()$  do
24        if  $D(v, q) < bsf$  then
25           $Knn.SortedInsert(v, D(v, q))$ ;
26    else
27      foreach  $N'$  in  $N.ChildNodes()$  do
28        if  $D_{lb}(N', q) < bsf$  then
29           $pq.Add(N')$ ;

// copy new results to global kNN Queue
30  $CopyResults(Knn, KnnResults, UpdateStart, UpdateEnd)$ ;
// Wait for other threads to finish current
   increment
31 Thread blocks on Barrier;
```

ABSTRACT

300 word description of the project

PVLDB Reference Format:

Jaouhara Chanchaf and Karima Echihabi. Paper Title, 15(10):
 XXX-XXX, 2022.
 doi:XX.XX/XXX.XX

1 PROPOSED APPROACH

Algorithm 1: KASHIF: INITTHREADPOOL

Input: $JobArray = \{q_1, \dots, q_n\}$, k , r_{th} , m the number of threads, $JobCounter$, $AllWorkersDone$.
Output: $KnnResults$.

```
1 Thread Array Workers[m];
2 Barrier Barrier;
3 Boolean AllWorkersDone  $\leftarrow$  False;
4 Shared Queue KnnResults[n][k]  $\leftarrow$ 
   $\{\{+\infty_1, \dots, +\infty_k\}_{q_1}, \dots, \{+\infty_1, \dots, +\infty_k\}_{q_n}\}$ ;
5 for  $i \leftarrow 0$  to  $m - 1$  do
6   Workers[i].INITWORKERTHREAD( $JobArray, k,$ 
7    $JobCounter, Barrier, KnnResults, AllWorkersDone$ );
  // get incremental results
8 while !AtomicLoad( $AllWorkersDone$ ) and do
9   Coordinator threads waits for new results on Barrier;
  if Recall( $KnnResults, GroundThru$ )  $\geq r_{th}$  then
10   StopAllWorkers();
11   Return KnnResults;
```

ACKNOWLEDGMENTS

We sincerely thank X, Y and Z.

REFERENCES

- [1] K. Echihabi, K. Zoumpatianos, T. Palpanas, and H. Benbrahim. The Lernaean Hydra of Data Series Similarity Search: An Experimental Evaluation of the State of the Art. *PVLDB*, 12(2), 2018.
- [2] PhDComics. Graduate Student Work Output. <https://phdcomics.com/comics/archive.php?comid=124>, 2022.

This work is licensed under the Creative Commons BY-NC-ND 4.0 International License. Visit <https://creativecommons.org/licenses/by-nc-nd/4.0/> to view a copy of this license. For any use beyond those covered by this license, obtain permission by emailing info@vldb.org. Copyright is held by the owner/author(s). Publication rights licensed to the VLDB Endowment.
Proceedings of the VLDB Endowment, Vol. 15, No. 10 ISSN 2150-8097.
doi:XX.XX/XXX.XX