The V*-Diagram: A Query-Dependent Approach to Moving KNN Queries

Sarana Nutanong, Rui Zhang, Egemen Tanin, Lars Kulik

Dept. of Computer Science and Software Engineering
University of Melbourne
Motivation

Consider two scenarios:

- a driver in a GPS-equipped car finding the nearest gas station along the route of a trip;
- an ambulance maintaining a list of \( k \) nearest hospitals while driving around a city.

These scenarios are examples of moving \( k \) nearest neighbor queries (\( MkNN \)).
Simple Approach
The Voronoi Diagram

Figure 1: Voronoi diagrams

Drawbacks:

1. Expensive precomputations (quadratic wrt $k$ [Aurenhammer and Schwarzkopf])
2. Inefficient update operations
3. No support for dynamically changing $k$ values
Best Existing Approach
Influence-set Retrieval [Zhang et al., 2003]

(a) Bisector $B_{ad}$ is discovered as a boundary.

(b) All boundaries are discovered

Figure 2: Computing a Voronoi cell locally
Our Approach: V*-Diagram

Objectives:
1. Requires \textit{no precomputation}
2. Supports \textit{insertions} and \textit{deletions} of objects
3. Handles \textit{dynamically changing} $k$

\begin{itemize}
  \item \textbf{W}($q_1$,f)
  \item $S(q_1,f,c)$
  \item $S(q_1,f,a)$
  \item \textbf{B}_{bc}
  \item \textbf{B}_{bf}
  \item \textbf{B}_{ac}
  \item \textbf{d}
  \item \textbf{e}
  \item \textbf{a}
  \item \textbf{b}
  \item \textbf{c}
  \item \textbf{f}
\end{itemize}
Our Approach: V*-Diagram

Objectives:

1. Requires *no precomputation*
2. Supports *insertions and deletions* of objects
3. Handles *dynamically changing k*

Result: Outperforms the best practice [Zhang et al.] by *2 orders of magnitude*
If the known NNs to $q$ are \( \{d, f, j\} \),
the known region $W(q, j)$ is
\[
\{v : \text{dist}(q, v) \leq \text{dist}(q, j)\}.
\]
The V*-Diagram
Safe region wrt a data point

We retrieve \((k + x)\) objects. In this example, \(k\) and \(x\) are 1, so we retrieve \(p\) and \(z\).

If \(q' \in S(q_b, z, p)\) then,

\[
\forall p' \not\in W(q_b, z), \dist(q', p) < \dist(q', p').
\]

\[
S(q_b, z, p) = \{ q' : \dist(p, q') \leq \dist(q_b, z) - \dist(q_b, q') \}.
\]
The V*-Diagram
The Fixed-rank Region (FRR) [Kulik and Tanin, 2006]

Figure 3: Incremental rank update
The V*-Diagram
Integrated Safe Region (ISR) and V*-kNN

ISR is an intersection of
1. the safe region wrt $k^{th}$ NN, $S(q_b, z, p_k)$;
2. the FRR of the $(k+x)$ NNs of $q_b$. 

Figure 4: V*-kNN Example ($k = 2$, $x = 2$)
V*-kNN Algorithm

Experiments

- Data Structure: R*-trees (1-kB block size)
- Comparative Method: RIS-$k$NN [Zhang et al.]
- Trajectories:

Figure 5: Trajectory types
Experiments

The two datasets are:

(a) 65,743 postal addresses from California
(b) 119,897 postal addresses from North-Eastern USA

Figure 6: Effect of $k$
Experiments
total cost wrt $n$

We use two types of distributions, uniform and Zipfian.

Figure 7: Effect of dataset size
Conclusions

- The V*-Diagram constructs a safe region using:
  1. the location of the query point,
  2. $k$NN-search coverage (known region),
  3. known data points.

- V*-kNN is *local*, *incremental* and *dynamic*.

- V*-kNN outperforms the best existing technique by two orders of magnitude.
Continuing Work
The V*-Diagram in a spatial network

Figure 8: The V*-Diagram in a spatial network ($k = 1$ and $x = 2$)
Acknowledgments

This work is supported under

• the Australian Research Council’s Discovery funding scheme (project number DP0880215);
• NICTA Victoria Laboratory.