

## REVERSE NEAREST NEIGHBOURS DENGAN REGION

Dimas Satrio Hutomo H.<sup>1</sup>, Warih Maharani<sup>2</sup>, Kiki Maulana<sup>3</sup>

<sup>1</sup>Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

---

### Abstrak

Reverse Nearest Neighbours (RNN) merupakan sebuah variasi dari konsep Nearest Neighbours (NN) yang dapat digunakan untuk menentukan tetangga terdekat dari suatu query point. RNN memiliki sifat inversal yang unik dimana RNN memberikan hasil yang berbeda dan tidak simetris dari konsep Nearest Neighbours biasa. Ini memberikan alternatif solusi terhadap pemrosesan dari Nearest Neighbours biasa sehingga layak untuk dipelajari lebih jauh. Namun komputasi RNN yang ada saat ini menghabiskan sumber daya yang sangat besar dan sangat tidak efisien. Kelemahan RNN saat ini menyebabkan RNN tidak dapat digunakan pada dataset bergerak dan dinamis. Sehingga tidak cocok untuk service berbasis lokasi yang sering digunakan pada perangkat mobile. Dengan memanfaatkan pendekatan menggunakan konsep region, kandidat RNN dapat dihasilkan tanpa melaksanakan pengecekan balik ke tiap titik yang ada pada dataset, seperti pada konsep RNN aslinya. Sebuah konsep Contact Zone yang akan dijelaskan pada jurnal ini dapat digunakan untuk mengurangi kandidat yang diproses sehingga beban komputasi dapat ditekan.

**Kata Kunci :** reverse nearest neighbours, influence zone, nearest neighbours, contact zone, peers, voronoi diagram.

---

### Abstract

Reverse Nearest Neighbour (RNN) is one of the variant of Nearest Neighbours (NN) algorithm dapat used to select nearest neighbour of a query point. RNN concept of inverse is capable to generate a different set of result and not symmetrical from a common nearest neighbours concept. Yet standard RNN computation spend a lot of reseed and highly inefficient. This weakness makes RNN not suitable for location based service that usually answered by common nearest neighbor. By utilizing an approach of region concept, RNN candidate can be generated without rechecking each point inside dataset. Contact Zone concept that will be explained in this journal can be used to eliminate processed candidate without processing all data in space, making it available for use with dynamic and mobile dataset system.

**Keywords :** reverse nearest neighbours, influence zone, nearest neighbours, contact zone, peers, voronoi diagram.

---

Telkom  
University

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

*k-Nearest Neighbour* (kNN) merupakan salah satu metode yang digunakan pada sistem informasi berbasis lokasi (*location based service*). kNN digunakan dalam pencarian objek tetangga terdekat dari sebuah titik di dalam ruang yang tidak dapat diproses dengan menggunakan algoritma pencarian lokasi biasa. Selain itu, kNN pun dapat digunakan untuk mencari suatu lokasi ketika *layout* dari lokasi tersebut tidak diketahui. Hal ini menjadikan kNN digunakan pada dataset yang bersifat dinamis dan dapat berubah setiap saat.

Salah satu variasi dari algoritma kNN adalah *Reverse Nearest Neighbour* (RNN). Kelebihan dari algoritma RNN adalah adanya sifat tidak simetris dengan kNN sehingga dapat memberikan variasi solusi bagi permasalahan yang hanya bias diselesaikan oleh kNN. Pada prosesnya, RNN memanfaatkan penggunaan *R-Tree* sebagai sarana pencarian suatu *query*. Namun hal ini menghalangi penggunaan *dataset* yang bersifat dinamis karena waktu dan komputasi yang dibutuhkan tinggi, sehingga dapat berakibat pada perubahan struktur data. <sup>[4][5]</sup>

Untuk mendapatkan hasil yang lebih baik dalam algoritma RNN, digunakan sebuah metode baru yang memanfaatkan komputasi geometri, yaitu pemanfaatan *contact zone*. Dengan *contact zone*, proses pengecekan titik-titik dapat dikurangi dengan cara hanya memeriksa titik yang berada di dalam *contact zone* <sup>[2]</sup>. Hal ini dapat mengeliminasi kebutuhan untuk menciptakan *R-Tree* dan dapat menggunakan *dataset* yang dinamis.

### 1.2. Perumusan Masalah

Pada tugas akhir ini, masalah yang dijadikan objek penelitian adalah sebagai berikut.

- a. Bagaimana mengatasi kelemahan dari algoritma RNN.
- b. Bagaimana nilai performansi penggunaan peers pada algoritma berbasis region dalam menyelesaikan masalah mengenai RNN.
- c. Bagaimana hasil implementasi sistem dari algoritma RNN dengan region pada dataset dengan jumlah  $> 100$  ?
- d. Bagaimana hasil pengujian dengan cara membandingkan kemiripan hasil dengan pengujian konseptual yang telah dilaksanakan sebelumnya?

### **1.3. Tujuan**

Tujuan dari penelitian tugas akhir ini adalah :

- a. Menggunakan *RNN* dengan pendekatan region sebagai solusi pemecahan masalah yang berkaitan dengan *influence zone* suatu *query* dalam sebuah *dataset* berbasis lokasi.
- b. Menilai performansi penggunaan peers pada algoritma berbasis region dalam penyelesaian masalah RNN.
- c. Implementasi algoritma RNN dengan region pada dataset dengan jumlah 100.
- d. Melakukan pengujian hasil implementasi algoritma RNN dengan region dengan cara membandingkan hasil implementasi dengan hasil pengujian konseptual.

### **1.4. Batasan Masalah**

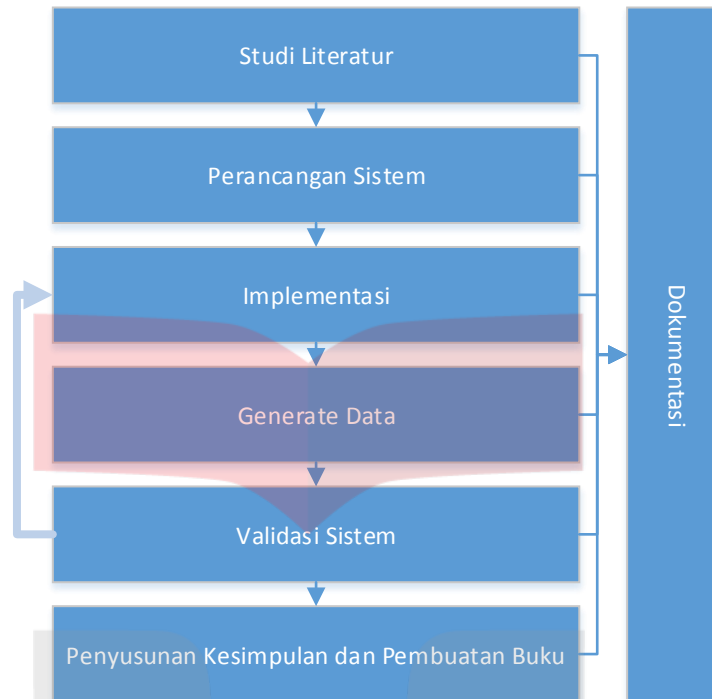
Adapun batasan masalah dalam penelitian tugas akhir ini adalah :

- a. Penggunaan *RNN* dilakukan pada *influence zone* dari sebuah *query point* dalam sebuah dataset yang berbasis lokasi.
- b. Optimasi region pembagi ruang dari solusi RNN hanya dilakukan untuk memperbaiki performansi dalam kasus tertentu.
- c. Penelitian hanya menganalisa RNN yang berjenis *bichromatic* (dengan parameter *bichromatic* ( $k$ ) sama dengan satu).

### **1.5. Metodologi Penyelesaian Masalah**

Dalam penelitian tugas akhir ini penulis menerapkan pendekatan sistematis untuk memecahkan masalah di atas. Adapun metode penelitian yang dilakukan adalah sebagai berikut :

Telkom  
University



Gambar 1-1 Metode penyelesaian masalah.

#### e. Studi Literatur

Pembelajaran melalui studi literatur dengan menggunakan bahan literatur yang mendukung dalam penelitian ini, dalam hal ini studi literatur yang digunakan oleh penulis antara lain :

- Literatur yang berhubungan dengan solusi *Nearest Neighbours*
- Literatur yang berhubungan dengan konteks *dynamic dataset* dalam masalah *influence zone* dari sebuah *query point*.
- Literatur yang berhubungan dengan komputasi geometri.
- Literatur yang berhubungan dengan materi *Reverse Nearest Neighbours*.
- Referensi lain yang membantu dalam penelitian tugas akhir ini.

#### f. Perancangan Sistem

Dalam merancang sistem yang digunakan untuk penelitian *Reverse Nearest Neighbours* dengan Region ini. Sistem harus mampu mendukung seluruh kasus yang dapat menjadi *input* bagi sistem. Selain itu sistem harus mampu memvisualisasikan *output* solusi dari *input*, untuk kepentingan validasi algoritma tersebut. Selain itu sistem dapat secara modular ditambahkan fitur-fitur yang dapat membantu penelitian. Dan sistem dapat melakukan proses otomatisasi untuk melakukan pengecekan atas data yang berjumlah banyak. Hasil dari sistem ini berupa visualisasi dari *influence* sebuah *query point*, beserta info lain yang relevan seperti banyaknya pers yang terlibat, waktu eksekusi program, *error event* bila terjadinya *error*.

g. Implementasi

Pada tahap ini sistem akan dibangun sesuai tahap pada perancangan sistem. Sistem akan menggunakan bahasa Java dengan berbasis *desktop*, dengan harapan dapat dijalankan pada seluruh perangkat yang menggunakan *Java Runtime Environment*.

h. *Generate* data yang digunakan.

Sistem akan menggunakan data yang telah di *generate* sebelumnya sebagai *input* untuk melaksanakan proses validasi dan pengujian. Secara aktual sistem harus mampu menangani seluruh jenis data yang digunakan beserta seluruh kasus yang dapat timbul saat eksekusi program. Data terdiri atas *point* pada bidang *xy* yang di *generate* dengan nilai random dan didistribusikan secara uniform pada bidang *xy*.

i. Validasi Sistem.

Pada tahap ini sistem diuji validitasnya dengan *input dataset* yang di *generate* sebelumnya, pengujian ini untuk mengukur pengaruh dari penerapan algoritma region pada sistem *RNN*. Adapun parameter yang menjadi patokan adalah jumlah *peers* yang diproses dalam eksekusi program. Selain itu *output* secara visual dapat digunakan untuk memvalidasi program secara matematis.

j. Penyusunan kesimpulan dan pembuatan buku.

Analisa akan disusun berdasarkan kesimpulan yang didapat dari pengujian program. Pengaruh dari penggunaan algoritma region dapat dilihat berdasarkan jumlah *peers* yang diproses maupun kecepatan eksekusi program. Dan dapat dibandingkan dengan algoritma lain yang menghasilkan hasil yang sama.

k. Dokumentasi

Dokumentasi dilakukan untuk melakukan setiap terjadinya *milestone* atau hal-hal penting lain yang terjadi saat pelaksanaan tugas akhir ini.

Telkom  
University

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1. Kesimpulan

Berdasarkan dari hasil pengujian dan analisis terhadap sistem yang dibangun dengan menggunakan metode Contact Zone dengan Initial Region dan Contact Zone tanpa Initial Region, dapat ditarik kesimpulan sebagai berikut.

- a) Metode Initial Region dengan menggunakan contact zone dapat diimplementasikan ke dalam dataset yang memiliki jumlah  $> 100$ . Hasil implementasi dapat memberikan influence zone yang dicari dengan performansi yang lebih baik daripada metode RNN yang melakukan pendekatan dengan cara pengecekan tiap titik.
- b) Implementasi metode Contact Zone tanpa Initial Region memiliki performansi yang lebih baik dalam menyelesaikan kasus pencarian RNN Region pada monochromatic RNN. Hal ini dikarenakan metode ini tidak memerlukan pemrosesan Initial Region saat eksekusinya sehingga memproses lebih sedikit peers daripada metode Contact Zone dengan Initial Region.
- c) Implementasi metode Contact Zone tanpa Initial Region pun lebih handal dan aman digunakan, hal ini karena metode ini secara baik mengeliminasi peers yang tidak akan mempengaruhi hasil region sejak awal. Eliminasi ini pun dapat segera dilakukan sejak vertex pertama terbentuk. Komputasi yang rendah serta processing time yang rendah cocok digunakan dalam aplikasi yang bekerja pada perangkat Mobile maupun aplikasi berbasis dataset dinamis lain.
- d) Initial Region hanya merupakan kandidat dari RNN Region yang dicari, hal ini karena dalam kebanyakan kasus, Initial Region mengalami reuksi ukuran akibat pemrosesan peers yang mengakibatkan terciptanya vertex baru.

#### 5.2. Saran

Setelah menyelesaikan proses pengerjaan tugas akhir ini, penulis memiliki beberapa saran sebagai berikut :

- a) Untuk pengembangan selanjutnya, akan sangat baik bila dataset yang diuji cobakan memiliki jumlah yang cukup banyak beserta variasi persebaran yang beraneka ragam.
- b) Bila suatu sistem hanya membutuhkan jawaban tanpa memperhatikan akurasi secara detail, suatu sistem yang menggabungkan kedua metode tersebut dapat dibangun. Dengan Initial Region, sistem dapat memberikan perkiraan RNN dengan lebih cepat. Sehingga dapat memberikan respon kepada user dengan lebih cepat. dan bilamana proses *peers* melebihi suatu nilai *epsilon* tertentu maka sistem ini dapat mengakhiri proses eksekusi, untuk selanjutnya diproses pada metode

Contact Zone tanpa Initial Region. Penulis juga menemukan bahwa kompleksitas program dapat dikurangi dengan membuang setiap bisector yang tidak menghasilkan vertex yang valid pada pemrosesan. Dengan membuang bisector ini, proses pengulangan dapat ditekan serendah mungkin dan mengurangi kompleksitas dari program.



## DAFTAR PUSTAKA

- [1] A. B. Waluyo, B. Srinivasan och D. Taniar., "Research in mobile database query optimization and processing. Mobile Information Systems.," 2005.
- [2] K. M. Adhinugaraha, D. Taniar och M. Indrawan, Finding Reverse Nearest Neighbours by Region, Wiley Interscience, 2013.
- [3] L. Bohan och Q. Xiaolin, "Research on Reverse Nearest Neighbor Using Ranked Voronoi Diagram," i *International Conference of Information Science and Engineering (ICISE2009)*, 2009.
- [4] R. Stanoi, D. Agrawal och A. E. Abbadi, Reverse Nearest Neighbour Queries for Dynamic Databases., 2000.
- [5] F. Korn och S. Muthukrishnan, "Influence Sets Based on Reverse Nearest Neighbor Queries.," *AT&T Research Laboratory.*, 2009.
- [6] M. A. Cheema, X. Lin, W. Zhang och Y. Zhang, "Influence Zone : Efficiently Processing Reverse k Nearest Neighbor."
- [7] M. Safat, D. Ibrahim och D. Taniar, "Voronoi based reverse nearest neighbor queries on spatial networks.," *Springer-Verlag*, 2009.
- [8] M. v. Kreveld, "Department of Information and Computing Sciences. Faculty of Science, Utrecht University.," May 2013. [Online]. Available: <http://www.cs.uu.nl/docs/vakken/ga/>. [Använd September 2013].
- [9] T. Switzer, "Tom Switzer 2D Convex Hulls: Jarvis March. The simple approach to computing the convex hull in 2D.," 12 2009. [Online]. Available: <http://tomswitzer.net/2009/12/jarvis-march/>. [Använd 7 2013].
- [10] A. Guttman, "R-TREES. A DYNAMIC INDEX STRUCTURE," University of California, 1984.