LAPORAN TUGAS KECIL 2

IF2211 STRATEGI ALGORITMA PENCARIAN TITIK TERDEKAT DENGAN ALGORITMA DIVIDE AND CONQUER

Disusun untuk memenuhi tugas mata kuliah Strategi Algoritma pada Semester 2 (dua) Tahun Akademik 2022/2023.



Oleh:

13521087 Razzan Daksana Yoni 13521095 Muhamad Aji Wibisono

PROGRAM STUDI TEKNIK INFORMATIKA SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA INSTITUT TEKNOLOGI BANDUNG 2022

DAFTAR ISI

DAFTAR ISI	2
BAB I	
PENDAHULUAN	3
1.1. Latar Belakang	3
BAB II	
ALGORITMA	4
2.1. Penjelasan Algoritma Divide and Conquer	4
2.2. Source Program dalam Bahasa Python	6
2.3. Analisis Algoritma	11
BAB III	
EKSPERIMEN	14
BAB IV	
KESIMPULAN	17
LAMPIRAN	18

BABI

PENDAHULUAN

1.1. Latar Belakang

Algoritma Divide and Conquer adalah algoritma yang menyelesaikan masalah dengan cara membagi permasalahan tersebut menjadi beberapa upa-persoalan yang memiliki kemiripan dengan persoalan semula namun berukuran lebih kecil (*divide*). Setelah dilakukan pembagian, algoritma akan menyelesaikan permasalahan jika sudah berukuran kecil atau kembali membagi secara rekursif jika masih berukuran besar (*conquer*). Di akhir algoritma, solusi masing - masing upa-persoalan akan digabungkan sehingga membentuk solusi permasalahan semula (*combine*).

Algoritma Divide and Conquer dapat menghasilkan solusi dengan kompleksitas yang lebih rendah daripada algoritma brute force dalam beberapa kasus sehingga dapat mempercepat pencarian solusi dalam skala tertentu.

Pada permasalahan pencarian titik terdekat, disajikan sejumlah titik acak dalam suatu ruang berdimensi tertentu. Lalu pada kumpulan titik acak tersebut dilakukan pencarian hingga dihasilkan pasangan titik dengan jarak satu sama lain paling dekat dibandingkan pasangan lainnya.

Pada laporan ini dipaparkan program yang telah dibuat untuk menyelesaikan permasalahan pencarian titik terdekat dengan algoritma Divide and Conquer. Selain itu pada program juga diperlihatkan perbandingan waktu serta kompleksitas penyelesaian masalah yang sama menggunakan algoritma Brute Force sebagai perbandingan. Rincian dari program yang dipaparkan terdapat pada bab - bab selanjutnya.

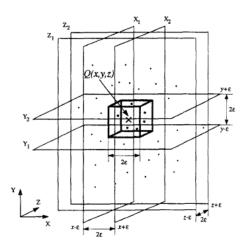
BAB II

ALGORITMA

2.1. Penjelasan Algoritma Divide and Conquer

Pada Algoritma Divide and Conquer yang digunakan ada beberapa langkah yang dilakukan:

- 1. Kumpulan titik diurutkan berdasarkan salah satu dimensinya.
- 2. Kumpulan titik akan dibagi menjadi dua ruas sama besar berurut berdasarkan salah satu dimensinya. *(divide)*
- 3. Jika jumlah titik adalah dua maka didapatkan pasangan titik terdekat, jika jumlah titik adalah tiga maka akan dilakukan pencarian secara Brute Force *(conquer)*. Jika lebih dari tiga maka akan dilakukan pembagian lagi sampai didapatkan kumpulan yang berjumlah dua atau tiga.
- 4. Setelah didapatkan jarak terdekat dari masing masing ruas akan diperhitungkan jarak antara titik titik yang berada pada perbatasan ruas satu dan yang lainnya dengan titik yang perbatasan didefinisikan dengan titik yang berjarak kurang dari jarak terdekat yang didapatkan dari kedua belah ruas. (combine)
- 5. Perhitungan pada titik titik yang berada pada perbatasan ruas satu dan yang lainnya dilakukan dengan melakukan *traversal* terhadap kumpulan lalu membagi titik menjadi dua belah array yang terdapat pada perbatasan untuk yang berada pada satu sisi dan yang lainnya.
- 6. Setelah terbentuk dua array yang berisikan titik yang berada pada perbatasan pada satu sisi dan yang lainnya, untuk setiap titik pada satu array akan dilakukan pengecekan untuk setiap titik pada array lainnya bahwa titik tersebut memenuhi kriteria pengecekan yaitu jarak pada semua dimensinya kurang dari jarak terdekat yang telah didapatkan sebelumnya sedangkan titik yang tidak memenuhi hanya dilewatkan saja. Ada pula *visualisasi* pada ruang tiga dimensi adalah sebagai berikut:



Sumber: Closest Point Search in High Dimensions - Computer Vision and Pattern Recognition, 1996. Proceedings CVPR '96, 1996 IEEE Computer Society Co (columbia.edu)

- 7. Jika jarak terdekat dari titik titik perbatasan lebih dekat dari yang didapat sebelumnya, maka jarak terdekat sekarang adalah jarak terdekat dari titik titik perbatasan.
- 8. Jarak terdekat akan terus dibandingkan sampai didapat kumpulan titik sejumlah semula. Setelah itu didapatlah jarak terdekat dari sekumpulan titik tersebut.

2.2. Source Program dalam Bahasa Python

Program dibuat dengan antarmuka teks. Selain dapat membuat kumpulan titik secara random, program juga dapat membaca kumpulan titik dari file .txt.

```
tool.py
      • • •
```

```
from time import time as currentTime
   from random import uniform as randomUniform
   def initializeCounter():
       euclidN = 0
8 def getDistanceBetweenPoints(point1:list, point2:list, dimension:int = 3) -> float:
       euclidN += 1
       for i in range(dimension):
16 def validchecker(point1: list, point2: list, threshold: float, startcheck: int = 0) -> bool:
       for i in range(startcheck, len(point1)):
24 def sortArraym(arrOfPoint:list, key:int, n:int):
           arrOfPoint1 = arrOfPoint[:n//2]
          sortArraym(arrOfPoint1, key, len1)
          sortArraym(arrOfPoint2, key, len2)
          idx1 = 0
           idx2 = 0
               if (arr0fPoint1[idx1][key] < arr0fPoint2[idx2][key]):</pre>
                   arrOfPoint[idx] = arrOfPoint1[idx1]
                   idx1 += 1
                   arrOfPoint[idx] = arrOfPoint2[idx2]
              idx += 1
           while idx2 < len2:
               idx2 += 1
56 def sortArrayq(arrOfPoint:list, key:int, i:int, j:int):
           k = partite2(arrOfPoint, key, i, j)
           sortArrayq(arrOfPoint, key, i, k-1)
           sortArrayq(arrOfPoint, key, k+1, j)
```

```
pivot = arrOfPoint[i][key]
         while arrOfPoint[p][key] < pivot:</pre>
         arrOfPoint[p], arrOfPoint[q] = arrOfPoint[q], arrOfPoint[p]
             while arrOfPoint[p][key] < pivot:</pre>
        arrOfPoint[p], arrOfPoint[q] = arrOfPoint[q], arrOfPoint[p]
arrOfPoint[i], arrOfPoint[q] = arrOfPoint[q], arrOfPoint[i]
39 #referensi internet
40 def partite2(arrOfPoint:list, key:int, i:int, j:int):
         arrOfPoint[p], arrOfPoint[j] = arrOfPoint[j], arrOfPoint[p]
53 def readFile(fileName:str) -> tuple :
        dimension:int
        with open(fileName, 'r') as f:
                  dimension = len(temp)
                  for num in temp:
                  arrOfPoints = arrOfPoints + [point]
         return (arrOfPoints, n, dimension)
```

```
visualization.py

from matplotlib.pyplot import figure, show

def three_dimensional_plotting(arrOfPoint:list, pointA:list, pointB:list, name:str = ""):
    fig = figure(num=name)

    # 3D plot from array of points
    ax = fig.add_subplot(111, projection='3d')
    for point in arrOfPoint:
        ax.scatter(point[0], point[1], point[2], color='blue')

ax.scatter(pointA[0], pointA[1], pointA[2], color='red')
    ax.scatter(pointB[0], pointB[1], pointB[2], color='red')
    show()

14
```

divideNConquer.py

```
from bruteForce import getClosestPairByBruteForce
from tools import getDistanceBetweenPoints, validchecker
def getClosestPairByDivideNConquer(arrOfPoint:list, n:int, dimension:int = 3) -> tuple :
        return getClosestPairByBruteForce(arrOfPoint, n, dimension)
        arrOfPoint2 = arrOfPoint[n // 2:]
        pointA1, pointB1, minDistance1 = getClosestPairByDivideNConquer(arrOfPoint1, n // 2, dimension) pointA2, pointB2, minDistance2 = getClosestPairByDivideNConquer(arrOfPoint2, n - (n // 2),
dimension)
else (pointA2, pointB2, minDistance2)
        nMiddleLine1 = 0
        arrOfPointInMiddleLine2 = []
        for i in range(n):
                 arrOfPointInMiddleLine1 = arrOfPointInMiddleLine1 + [arrOfPoint[i]]
                arrOfPointInMiddleLine2 = arrOfPointInMiddleLine2 + [arrOfPoint[i]]
                 if validchecker(arr0fPointInMiddleLine1[i], arr0fPointInMiddleLine2[j], minDistance, 1):
                     distance = getDistanceBetweenPoints(arrOfPointInMiddleLine1[i],
arrOfPointInMiddleLine2[j], dimension)
                     if distance < minDistance:</pre>
                         minDistance = distance
            # nPointToCheck = nMiddleLine
             for j in range(1, nPointToCheck+1):
                 if abs(arrOfPointInMiddleLine[i][1] - arrOfPointInMiddleLine[i+j][1]) < minDistance :</pre>
                     distance = getDistanceBetweenPoints(arrOfPointInMiddleLine[i], arrOfPointInMiddleLine
[i+j], dimension)
                      if distance < minDistance:</pre>
                          pointA = arrOfPointInMiddleLine[i]
                          pointB = arrOfPointInMiddleLine[i+j]'''
        return (pointA, pointB, minDistance)
```

main.py

```
• • •
    from bruteForce import getClosestPairByBruteForce
   from divideNConquer import getClosestPairByDivideNConquer from visualization import three_dimensional_plotting
    def main():
        setrecursionlimit(100000)
             print("""
13 Input action:
14 1. Random
   2. Read File
                  inputUser = int(input("-> "))
                  if inputUser == 1 :
                      n = int(input("Enter number of points: "))
                         raise Exception("Number of points must be larger than 1")
                       dimension = int(input("Enter dimension: "))
                           raise Exception("Dimension must be larger than 0")
                           temp = [tools.randomUniform(-100, 100) for j in range(dimension)]
                       fileName = input("Input File Name: ")
                  print(exception)
             inputUser = -1
if ((inputUser == 1) or (inputUser == 2)) :
                  tools.initializeCounter()
                  timeBruteForce = tools.currentTime()
                  point1BF, point2BF, minDistanceBF = getClosestPairByBruteForce(arrOfPoint, n,
    dimension)
                  print("Closest pair is Using Brute Force: ")
                  print("Point A: ", point1BF)
print("Point B: ", point2BF)
print("Distance: ", minDistanceBF)
                  print("Time: ", timeBruteForce)
                  print("Times euclidean distance called: ", tools.euclidN)
                  tools.initializeCounter()
                  tools.sortArrayq(arrOfPoint, 0, 0, n-1)
                  point1DC, point2DC, minDistanceDC= getClosestPairByDivideNConquer(arrOfPoint, n,
                  print("Closest pair is Using Divide and Conquer: ")
                  print("Point A: ", point1DC)
print("Point B: ", point2DC)
print("Distance: ", minDistanceDC)
                  print("Time: ", timeDivideNConquer)
                  print()
                       three_dimensional_plotting(arrOfPoint, point1BF, point2BF, "Brute Force")
three_dimensional_plotting(arrOfPoint, point1DC, point2DC, "Divide and Conquer")
         name
         main()
```

2.3. Analisis Algoritma

Algoritma Brute Force memiliki kompleksitas diperhitungkan secara sederhana dengan jumlah *loop* yang terdapat yaitu O(n²) untuk semua kasus yaitu dengan cara memilih semua pasangan titik lalu memanggil perhitungan *euclidean distance* dalam tiap *loop*.

Di sisi lain, Algoritma Divide and Conquer memiliki kompleksitas yang tidak sesederhana algoritma Brute Force. Merujuk kepada *source code* yang tertera pada bagian 2.2, algoritma Divide and Conquer memerlukan kumpulan titik untuk diurutkan terlebih dahulu sebelum dilakukan Divide and Conquer. Pengurutan yang diimplementasikan, Quick Sort, memiliki kompleksitas $O(n \log n)$. Setelah itu, algoritma Divide and Conquer memiliki rekursif yang memanggil dirinya sendiri dengan jumlah n dibagi kurang lebih sama rata menjadi dua bagian sehingga menghasilkan $T(\frac{n}{2})$. Untuk n yang bernilai 2, maka kompleksitas adalah 1 dan untuk n yang bernilai 3 kompleksitas adalah 3, ini menghasilkan notasi O(1). Saat penggabungan dilakukan traversal kembali untuk kedua kumpulan titik yang berada pada perbatasan. Pada kemungkinan terburuk, semua titik berada pada perbatasan sehingga pemilihan yang dilakukan adalah $(\frac{n}{2})^2$ didapatkan notasi yaitu $O(n^2)$. Secara kasus terburuk, algoritma Divide and Conquer memiliki kompleksitas:

$$T(n) = 1, n = 2$$

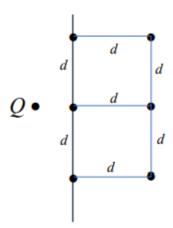
$$T(n) = 3, n = 3$$

$$T(n) = 2(T(n/2)) + O(n^2), n > 3$$

Menggunakan teorema master dengan a = 2, b = 2, dan d = 2, didapat kompleksitas pada algoritma Divide and Conquer setelah ditambahkan kompleksitas pengurutan adalah $O(n^2 + n \log n) = O(n^2)$.

Akan tetapi, perhitungan kompleksitas yang dilakukan pada algoritma Divide and Conquer sebelumnya berbeda dengan perhitungan kompleksitas yang dilakukan pada algoritma Brute Force. Perbedaan terbesar adalah bahwa pada algoritma Brute Force, kompleksitas algoritma berlaku untuk perhitungan *euclidean distance*, sedangkan pada algoritma Divide and Conquer kompleksitas hanya berlaku untuk pemilihan titik. Perlu diketahui bahwa perhitungan *euclidean distance* memakai *resource* yang jauh lebih banyak daripada pemilihan titik karena perhitungan *euclidean distance* dilakukan dengan perkalian untuk setiap dimensi diikuti dengan pertambahan dan pengakaran sedangkan pemilihan titik hanya menggunakan perbandingan untuk setiap dimensi.

Jika dilakukan perhitungan ulang kompleksitas menggunakan jumlah pemanggilan *euclidean distance* saja, algoritma Divide and Conquer akan memiliki kompleksitas yang berbeda. Sorting menjadi tidak diperhitungkan karena tidak melakukan perhitungan *euclidean distance*. Perhitungan rekursif yang dilakukan selanjutnya memiliki kompleksitas yang sama seperti sebelumnya. Perbedaan besar muncul ketika dilakukan penggabungan antar ruas. Sebab jarak minimum antara dua titik pada kedua belah sisi sudah diketahui sebelumnya, dapat ditentukan jumlah maksimum titik yang berada pada cakupan pengecekan. Ada pula visualisasi pada dimensi ruang dua dimensi adalah sebagai berikut:



Sumber: Algoritma Divide and Conquer (itb.ac.id)

Dengan melakukan *scaling-up* pada dimensi yang lebih tinggi dan induksi secara matematis dapat ditentukan bahwa jumlah titik maksimum yang mungkin memenuhi kriteria pengecekan mengikuti persamaan $2 \times 3^{d-1}$ dengan d adalah jumlah dimensi ruang titik - titik tersebut berada. Dari perhitungan tersebut dapat terlihat bahwa perhitungan *euclidean distance* pada algoritma Divide and Conquer bukan dipengaruhi oleh n atau jumlah titik, melainkan jumlah dimensi ruang titik - titik tersebut berada. Dengan demikian, kompleksitas dari perhitungan *euclidean distance* titik - titik yang berada pada perbatasan pembagian kedua ruas untuk kemungkinan terburuk adalah $(\frac{n}{2}) \times (2 \times 3^{d-1})$, sehingga didapat notasi Big O adalah O(n).

Secara kasus terburuk, algoritma Divide and Conquer memiliki kompleksitas untuk perhitungan *euclidean distance*:

$$T(n) = 1, n = 2$$

$$T(n) = 3, n = 3$$

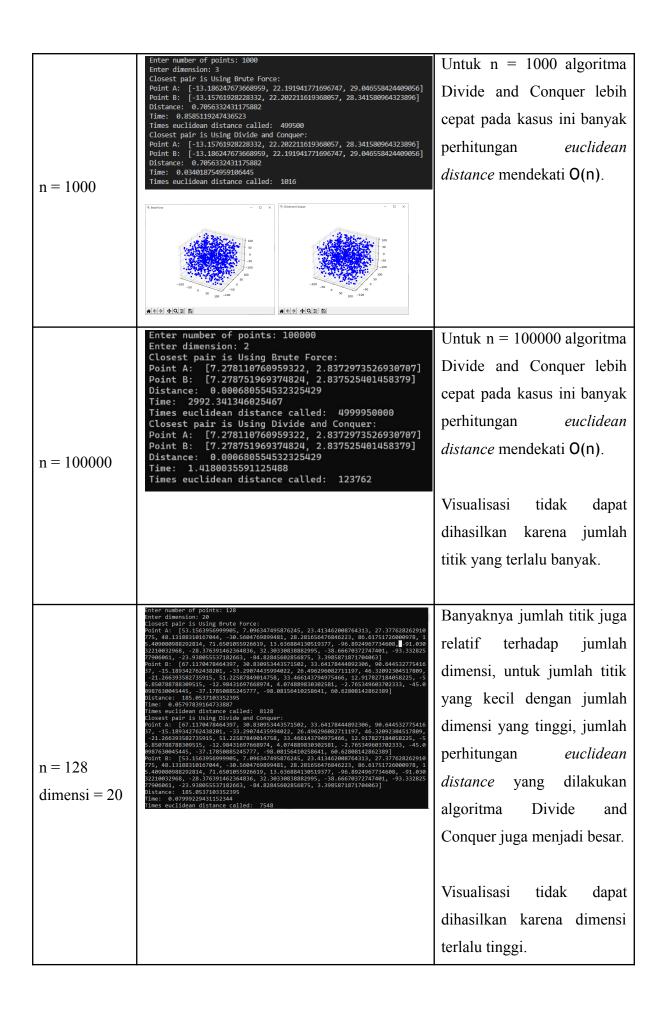
$$T(n) = 2(T(n/2)) + O(n), n > 3$$

Menggunakan teorema master dengan a = 2, b = 2, dan d = 1, didapat kompleksitas untuk kemungkinan terburuk perhitungan *euclidean distance* pada algoritma Divide and Conquer adalah $O(n \log n)$. Walaupun begitu jika perhitungan *euclidean distance* antara kedua ruas tidak akan melebihi jumlah titik secara total dan tidak akan perhitungan tidak akan dilakukan melebihi perhitungan yang dilakukan oleh Brute Force.

Dari analisis yang telah dilakukan, dapat disimpulkan bahwa algoritma Divide and Conquer relatif lebih sangkil dibandingkan algoritma Brute Force untuk jumlah titik yang tinggi dengan kompleksitas algoritma Divide and Conquer adalah O(n log n) dan kompleksitas algoritma Brute Force adalah $O(n^2)$. Hal ini dibuktikan dengan eksperimen yang telah dilakukan pada bab selanjutnya.

BAB III EKSPERIMEN

Test Case	Gambar	Keterangan	
n = 16	Enter number of points: 16 Enter dimension: 3 Closest pair is Using Brute Force: Point A: [-12.73394377/291855, 32.44886411571915, -98.55656510964073] Point B: [-24.99325172270295, 28.078090139415792, -97.30978303442244] Distance: 13.074470142399552 Time: 0.0 Times euclidean distance called: 120 Closest pair is Using Divide and Conquer: Point A: [-24.99325172270295, 28.078899139415792, -97.30978303442244] Point B: [-12.7339437772910855, 32.44886411571915, -98.55656510964073] Distance: 13.074470142399552 Time: 0.0 Times euclidean distance called: 14	Untuk n = 16 algoritma Divide and Conquer lebih cepat pada kasus ini banyak perhitungan euclidean distance mendekati O(n).	
n = 64	Enter number of points: 64 Enter dimension: 3 Closest pair is Using Brute Force: Point A: [35.90967874293787, 53.54417693019164, 91.62801341022376] Point B: [41.03182081261835, 56.88348457361835, 88.04738620613264] Distance: 7.085774911335809 Time: 0.00400090217590332 Times euclidean distance called: 2016 Closest pair is Using Divide and Conquer: Point A: [41.03182081261835, 56.88348457361835, 88.04738620613264] Point B: [35.90967874293787, 53.54417693019164, 91.62801341022376] Distance: 7.085774911335809 Time: 0.001001596450805664 Times euclidean distance called: 67	Untuk n = 64 algoritma Divide and Conquer lebih cepat pada kasus ini banyak perhitungan euclidean distance mendekati O(n).	
n = 128	Enter number of points: 128 Enter dimension: 3 Closest pair is Using Brute Force: Point A: [84.3078161068334, 28.365499291675974, -64.83391304625317] Point B: [83.67399344548647, 24.603186531327808, -64.82322039137594] Distance: 3.815343078359253 Time: 0.013014078140258789 Times euclidean distance called: 8128 Closest pair is Using Divide and Conquer: Point A: [83.67399344548647, 24.603186531327808, -64.82322039137594] Point B: [84.3078161068334, 28.365499291675974, -64.83391304625317] Distance: 3.815343078359253 Time: 0.003019571304321289 Times euclidean distance called: 123	Untuk n = 128 algoritma Divide and Conquer lebih cepat pada kasus ini banyak perhitungan euclidean distance mendekati O(n).	



 Jumlah dimensi juga mempengaruhi perhitungan karena relatif jumlah terhadap dimensi seperti telah disebutkan dalam sub bab sebelumnya.

Visualisasi tidak dapat dihasilkan karena dimensi terlalu tinggi.

BAB IV

KESIMPULAN

Pencarian dua titik terdekat dalam Rⁿ dapat diselesaikan dengan algoritma Brute Force dan algoritma Divide and Conquer. Dengan mengimplementasikan algoritma Divide and Conquer yang penulis tawarkan didapatkan jumlah perhitungan *euclidean distance* lebih sedikit untuk dimensi rendah dan jumlah titik banyak dibandingkan algoritma Brute Force. Dan didapatkan juga jumlah titik relatif terhadap jumlah dimensi yaitu jumlah titik yang kecil dengan jumlah dimensi yang tinggi membuat jumlah perhitungan *euclidean distance* pada algoritma Divide and Conquer juga menjadi besar.

LAMPIRAN

Link Spesifikasi Tugas

 $\underline{https://informatika.stei.itb.ac.id/}{\sim} \underline{rinaldi.munir/Stmik/2022-2023/Tucil2-Stima-2023.pdf}$

Cek List Spesifikasi

Poin	Ya	Tidak
Program berhasil dikompilasi tanpa kesalahan	1	
2. Program berhasil <i>running</i>	1	
3. Program dapat membaca menerima masukan dan menuliskan	1	
luaran		
4. Luaran program sudah benar (solusi <i>closest pair</i> benar)	1	
5. Bonus 1 dikerjakan	1	
6. Bonus 2 dikerjakan	1	

Repository Github

Berikut adalah link repository GitHub untuk program penulis.

https://github.com/razzanYoni/Tucil2_13521087_13521095