

**PERBANDINGAN WAKTU EKSEKUSI DAN PENGGUNAAN
MEMORI PADA PAKET GRAFIK PYTHON
UNTUK ANALISIS JARINGAN**



TESIS

Oleh:

WILDAN NADIYAL AHSAN
NIM. 20206052011
STATE UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PROGRAM STUDI INFORMATIKA
PROGRAM MAGISTER FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA

2024

PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

Nama : Wildan Nadiyah Ahsan
NIM : 20206052011
Program Studi : Informatika (S2)
Fakultas : Sains dan Teknologi

Menyatakan bahwa naskah tesis yang berjudul “Perbandingan Waktu Eksekusi dan Penggunaan Memori pada Paket Grafik Python untuk Analisis Jaringan” secara keseluruhan adalah hasil penelitian atau karya tulis saya sendiri, kecuali pada bagian-bagian tertentu yang dirujuk referensi sumbernya.

Yogyakarta, 18 Januari 2024
Saya yang menyatakan,



Wildan Nadiyah Ahsan
NIM. 20206052011

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PERNYATAAN BEBAS PLAGIASI

Yang bertanda tangan di bawah ini:

Nama : Wildan Nadiyah Ahsan
NIM : 20206052011
Program Studi : Informatika (S2)
Fakultas : Sains dan Teknologi

Menyatakan bahwa naskah tesis iyang berjudul “Perbandingan Waktu Eksekusi dan Penggunaan Memori pada Paket Grafik Python untuk Analisis Jaringan” secara keseluruhan benar-benar bebas dari plagiasi. Jika di kemudian hari terbukti melakukan plagiasi, maka saya siap ditindak sesuai ketentuan yang berlaku di Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga Yogyakarta.

Yogyakarta, 18 Januari 2024
Saya yang menyatakan,



Wildan Nadiyah Ahsan
NIM. 20206052011

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA



PENGESAHAN TUGAS AKHIR

Nomor : B-205/Un.02/DST/PP.00.9/01/2024

Tugas Akhir dengan judul : Perbandingan Waktu Eksekusi dan Penggunaan Memori pada Paket Grafik Python untuk Analisis Jaringan

yang dipersiapkan dan disusun oleh:

Nama : WILDAN NADIYAL AHSAN, S.Kom.
Nomor Induk Mahasiswa : 20206052011
Telah diujikan pada : Jumat, 26 Januari 2024
Nilai ujian Tugas Akhir : A-

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang

Dr. Agung Fatwanto, S.Si., M.Kom.
SIGNED

Valid ID: 65b87bd994466



Penguji I

Ir. Muhammad Taufiq Nuruzzaman, S.T.
M.Eng., Ph.D.
SIGNED

Valid ID: 65b86cd63a9da



Penguji II

Dr. Ir. Bambang Sugiantoro, S.Si., M.T.,
IPM., ASEAN Eng.
SIGNED

Valid ID: 65b75144af6dc



Yogyakarta, 26 Januari 2024

UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

Prof. Dr. Dra. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 65b8a3dba2453

NOTA DINAS PEMBIMBING

Hal: Persetujuan Tugas Akhir

Kepada Yth.,
Dekan Fakultas Sains dan Teknologi
UIN Sunan Kalijaga
Yogyakarta

Assalamu'alaikum wr.wb.

Setelah melakukan bimbingan, arahan, dan koreksi terhadap penulisan tesis yang berjudul:

**PERBANDINGAN WAKTU EKSEKUSI DAN PENGGUNAAN MEMORI PADA
PAKET GRAFIK PYTHON UNTUK ANALISIS**

Yang ditulis oleh:

Nama : Wildan Nadiyah Ahsan
NIM : 20206052011
Program Studi : Informatika (S2)
Fakultas : Sains dan Teknologi

Saya berpendapat bahwa tesis tersebut sudah dapat diajukan kepada Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga Yogyakarta untuk diujikan dalam rangka memperoleh gelar Magister Informatika.

Wassalamu'alaikum wr.wb.

Yogyakarta, 18 Januari 2024
Pembimbing,



Dr. Agung Fatwanto, S.Si., M.Kom.
NIP. 19770103 200501 1 003

ABSTRAK

Analisis data berbasis grafik adalah metode umum yang digunakan di berbagai domain. Beberapa paket grafik yang populer digunakan seperti NetworkX, igraph, NetworKit, dan graph-tool adalah contoh paket yang tersedia untuk bahasa pemrograman Python. Masing-masing memiliki kelebihan dan kekurangan. Penelitian ini bertujuan untuk membandingkan performa (dalam hal waktu eksekusi) dan penggunaan memori dari beberapa paket graf yang populer, yaitu NetworkX, Networkkit, SNAP, graph-tool, igraph, dan rustworkX. Penelitian ini menggunakan metode kuantitatif eksperimental. Dataset yang digunakan untuk simulasi analisis jaringan dengan menggunakan co-occurrence network adalah 10.000 judul artikel yang diambil dari beberapa jurnal yang terindeks scopus. Hasil penelitian menunjukkan bahwa nilai rata-rata \pm margin of error (pada tingkat kepercayaan 0,95 dan jumlah sampel/jumlah iterasi percobaan (n): 35) untuk waktu eksekusi tool graph-tool = $1.30291731 \pm 0.268629749$ detik (s), networkX = $2.07304831 \pm 0.672788959$ (s), networkkit = $1.44051549 \pm 0.272388017$ (s), snap = $0.92104769 \pm 0.126535737$ (s), igraph = $1.63668983 \pm 0.621590572$ (s), rustworkx = $1.33403117 \pm 0.190636021$ (s). Sementara itu, rata-rata \pm margin kesalahan untuk penggunaan memori saat ini dari graph-tool = $2.023255971 \pm 0.115873399$ MegaByte (MB) networkX = $5.1896566 \pm 2.431315437$ (MB), networkkit = $6.207405857 \pm 4.225013008$ (MB), snap = $3.743473629 \pm 1.080429925$ (MB), igraph = $4.461298114 \pm 1.20858089$ (MB), rustworkx = $4.267859943 \pm \dots$ (MB); dan rata-rata \pm margin kesalahan untuk penggunaan memori puncak dari graph-tool = $3.402223457 \pm 0.017638869$ MegaByte (MB), networkX = $5.630388457 \pm 2.813189597$ (MB), networkkit = $7.048138886 \pm 4.996092409$ (MB), snap = $4.120006486 \pm 1.288241414$ (MB), igraph = $4.825449429 \pm 1.47061189$ (MB), rustworkx = $4.621915486 \pm 1.130673792$ (MB). Hasil uji repeated measures ANOVA dan Friedman untuk waktu eksekusi, penggunaan memori saat ini, dan penggunaan memori puncak menunjukkan bahwa terdapat perbedaan yang signifikan di antara nilai rata-rata tersebut. Oleh karena itu, dapat disimpulkan bahwa urutan kinerja berdasarkan waktu eksekusi (dari yang tercepat ke yang paling lambat) adalah: 1. snap 2. graph-tool & rustworkx 3. networkkit & igraph 4. networkX. Sementara itu, urutan efisiensi berdasarkan penggunaan memori saat ini dan penggunaan memori puncak (dari yang paling efisien hingga yang paling tidak efisien) adalah: 1. graph-tool 2. snap 3. rustworkX 4. igraph 5. networkX 6. networkkit. Terdapat korelasi yang sedang dan positif antara waktu eksekusi dan penggunaan memori. Lebih jauh lagi, terdapat korelasi yang sangat kuat dan positif antara memori saat ini dan penggunaan memori puncak.

Kata Kunci: *Grafik Python; Komparasi Grafik Python.*

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

ABSTRACT

Graph-based data analysis is a prevalent method employed across various domains. Some popularly used graph packages such as NetworkX, igraph, NetworKit and graph-tool are examples of packages available for the Python programming language. Each has its advantages and disadvantages. This research aims to compare the performance (in terms of execution time) and memory usage of several popular graph packages, namely NetworkX, Networkkit, SNAP, graph-tool, igraph and rustworkX. The research employed an experimental quantitative method. The datasets used for simulating the network analysis by applying co-occurrence network was 10.000 article titles gathered from several journals indexed by scopus. The results shown that the mean value \pm margin of error (at level of confidence: 0.95 and number of samples/number of experiment iterations (n): 35) for the execution time of graph-tool = $1.30291731 \pm 0.268629749$ second (s), networkX = $2.07304831 \pm 0.672788959$ (s), networkkit = $1.44051549 \pm 0.272388017$ (s), snap = $0.92104769 \pm 0.126535737$ (s), igraph = $1.63668983 \pm 0.621590572$ (s), rustworkx = $1.33403117 \pm 0.190636021$ (s). Meanwhile, the mean \pm margin of error for the current memory usage of graph-tool = $2.023255971 \pm 0.115873399$ MegaByte (MB) networkX = $5.1896566 \pm 2.431315437$ (MB), networkkit = $6.207405857 \pm 4.225013008$ (MB), snap = $3.743473629 \pm 1.080429925$ (MB), igraph = $4.461298114 \pm 1.20858089$ (MB), rustworkx = $4.267859943 \pm \dots$ (MB); and the mean \pm margin of error for peak memory usage of graph-tool = $3.402223457 \pm 0.017638869$ MegaByte (MB), networkX = $5.630388457 \pm 2.813189597$ (MB), networkkit = $7.048138886 \pm 4.996092409$ (MB), snap = $4.120006486 \pm 1.288241414$ (MB), igraph = $4.825449429 \pm 1.47061189$ (MB), rustworkx = $4.621915486 \pm 1.130673792$ (MB). The results of repeated measures ANOVA and Friedman test for the execution time, current memory usage and peak memory usage shown that there is significant difference among these mean values. Therefore, it can be concluded that the order of the performance based on execution time (from fastest to slowest) is: 1.snap 2.graph-tool & rustworkx 3.networkkit & igraph 4.networkX. Meanwhile, the order of the efficiency based on the current memory usage and peak memory usage (from most efficient to most inefficient) is: 1.graph-tool 2.snap 3. rustworkX 4.igraph 5.networkX 6.networkkit. There was a moderate and positive correlation between execution time and memory usage. Further, there was a very strong and positive between current memory and peak memory usage.

Keywords: Python Graphs; Graphs Comparison.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

MOTTO

يَا إِلَهِي وَمَلِيكِي أَنْتَ تَعَلَّمَ كَيْفَ حَالِي



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PERSEMBAHAN

Puji syukur pada Tuhan Semesta Alam dan Shalawat teriring salam
teruntuk kekasih-Nya

Pada karya yang sederhana ini, saya persembahkan untuk:

Ibu saya tercinta Satuf Rohul Hidayah, S.E. dan ayah saya Alm. Dr. Saefudin Zuhri, M.A.,
tidak lupa pula kepada saudara kandung saya Tika Ifrida Takayasa, M.A., Adi Fikki
Al-Musta'an, S.H., dan Adik saya Bima Ahyadin Muhammad yang telah begitu
berjasa dalam memotivasi saya untuk terus tumbuh dan berkembang.

Seluruh Dosen dan Civitas Akademik Fakultas Sains dan Teknologi yang berjasa
selama menempuh pendidikan

Almamater Magister Informatika

Fakultas Sains dan Teknologi

Universitas Islam Negeri Sunan Kalijaga Yogyakarta

SUNAN KALIJAGA
YOGYAKARTA

TERIMAKASIH

KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

الحمد لله رب العالمين • وبه نستعين على امور الدنيا والدين • اشهد ان لا اله الا الله وحده لا شريك له وأشهد أن
محمدًا عبده ورسوله • اللهم صل على سيدنا محمد وعلى اله وصحبه اجمعين

Puji dan syukur penulis ucapkan kepada Allah swt. yang telah melimpahkan karunia-Nya berupa ilmu pengetahuan, kesehatan dan petunjuk sehingga tesis dengan judul “Perbandingan Waktu Eksekusi Dan Penggunaan Memori Pada Paket Grafik Python Untuk Analisis Jaringan” dapat diselesaikan dengan baik. Shalawat serta salam semoga selalu tercurahkan kepada Nabi Muhammad saw. dan segenap keluarga dan para sahabatnya yang tak kenal lelah memperjuangkan agama Islam yang ditunggu syafaatnya di *yaumil qiyāmah*.

Penulisan tesis ini dalam rangka memenuhi salah satu persyaratan untuk mencapai gelar Magister di Program Studi Informatika Universitas Islam Negeri Sunan Kalijaga Yogyakarta. Dalam menyelesaikan tesis ini, penulis menyadari dengan sepenuhnya bahwa terdapat banyak pihak yang turut serta membantu dalam proses penulisan tesis ini. Untuk itu, kepada seluruh pihak yang selama ini telah banyak membantu baik moril maupun materil, penulis mengucapkan terima kasih yang sebesar-besarnya. Ucapan terima kasih secara khusus penulis sampaikan kepada:

1. Prof. Dr. Phil Al Makin, S.Ag., M.A, selaku Rektor Universitas Islam Negeri Sunan Kalijaga Yogyakarta;
2. Dr. Dra Hj. Khurul Wardati, M. Si, selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga Yogyakarta;
3. Dr. Ir. Bambang Sugiantoro, S.Si., M.T, selaku Kepala Prodi Magister Informatika Universitas Islam Negeri Sunan Kalijaga Yogyakarta;
4. Dr. Ir. Sumarsono, S.T., M.Kom, selaku Sekretaris Prodi Magister Informatika Universitas Islam Negeri Sunan Kalijaga Yogyakarta;
5. Dr. Agung Fatwanto, S.Si., M.Kom, selaku pembimbing yang telah berkenan merelakan waktu, tenaga, dan ilmunya guna memberikan bimbingan kepada penulis dalam

menyelesaikan tesis ini, serta ucapan terimakasih dan penghargaan yang setinggi-tingginya, yang dengan penuh kesabaran dan kearifan telah memberikan bimbingan, arahan, dan dorongan di sela-sela kesibukannya;

6. KH. Muhammad Munawwar Ahmad, selaku pengasuh pondok pesantren Al-Munawwir Komplek L Kranyak yang telah memberikan barokah doa dan nasihat yang bermanfaat.
7. Dr. Bambang Sugiantoro, S.Si., MT. Selaku Dosen Penasehat Akademik yang telah berkenan membimbing dari proses awal perkuliahan sampai saat ini;
8. Seluruh Dosen Program Studi Magister Informatika Universitas Islam Negeri Sunan Kalijaga Yogyakarta yang telah membekali penulis dengan berbagai ilmu pengetahuan sehingga penulis dapat menyelesaikan tesis ini dengan baik;
9. Seluruh Staff Administrasi Program Studi Magister Informatika Universitas Islam Negeri Sunan Kalijaga Yogyakarta, yang secara langsung atau tidak langsung telah memberi bantuan kepada penulis dalam menyelesaikan penulisan tesis ini;
10. Yus Afrida, S.H., M.H. yang paling cantik, karena telah menemani, membantu serta kebersamai penulis selama proses penelitian berlangsung.
11. Bjong Cafe yang telah menyediakan tempat yang nyaman untuk menyelesaikan thesis.
12. Teman-teman seangkatan Program Studi Magister Informatika Universitas Islam Negeri Sunan Kalijaga Yogyakarta yang telah memberikan dukungan dalam penulisan tesis ini sampai selesai;
13. Ucapan terima kasih juga saya sampaikan kepada semua pihak yang tidak mungkin saya sebutkan satu demi satu yang telah banyak memberikan bantuan dan dukungan selama penyusunan tesis ini.

Semoga Allah swt. memberikan balasan yang berlipat ganda kepada semuanya. Penulis sadari masih banyak kekurangan dalam penulisan tesis ini, jika ada saran dan kritik yang membangun akan penulis terima dengan senang hati. Akhir kata, penulis berharap semoga tesis ini dapat bermanfaat khususnya bagi penulis dan umumnya bagi kita semua. *Aamiin...*

Yogyakarta, 18 Januari 2024

Saya yang menyatakan,



Wildan Nadiyah Ahsan
NIM. 20206052011



DAFTAR ISI

JUDUL	i
PPERNYATAAN KEASLIAN	ii
PERNYATAAN BEBAS PLAGIASI	iii
PENGESAHAN TUGAS AKHIR	iv
NOTA DINAS PEMBIMBING	v
ABSTRAK	vi
ABSTRACT	vii
MOTTO	viii
PERSEMBAHAN	ix
KATA PENGANTAR	x
DAFTAR ISI	xiii
DAFTAR TABEL	xvi
DAFTAR GAMBAR	xvii
DAFTAR BAGAN	xviii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	
A. Latar Belakang.....	1
B. Rumusan Masalah.....	4
C. Batasan Masalah.....	4
D. Tujuan penelitian.....	4
E. Manfaat Penelitian.....	5
F. Sistematika Penelitian.....	5
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI	
A. Tinjauan Pustaka.....	7
B. Landasan Teori.....	10
1. Graf.....	10
2. Analisis Jaringan Sosial (<i>Social Network Analysis</i>).....	13
3. Analisis Jaringan Teks (<i>Text Network Analysis</i>).....	15
4. Python Graph Package.....	17
a. NetworkX.....	17
b. Networkit.....	18

c. Igraph.....	18
d. Graph-tool.....	19
e. SNAP.....	20
f. RustworkX.....	20
5. Waktu Eksekusi (<i>Execution Time</i>).....	21
6. Penggunaan Memori (<i>Memory Usage</i>).....	22
7. Nilai Rata-Rata (MEAN).....	23
a. Mean Aritmatik.....	23
b. Mean Geometrik.....	24
8. Ukuran Sebaran.....	24
a. Standar Deviasi.....	24
b. Varians.....	25
c. Kovarians.....	25
9. Uji Normalitas.....	25
10. Uji Wilcoxon.....	27
11. Box Plot.....	27

BAB III METODE PENELITIAN

A. Desain Penelitian.....	29
1. Alur Penelitian.....	30
2. Studi Literatur.....	31
3. Dataset.....	32
4. Implementasi Analisis <i>Co-Occurance</i>	32
5. Proses Pengujian Performa.....	35
6. Analisis Data.....	39
B. Alat Penelitian.....	40
C. Pengambilan Kesimpulan.....	41

BAB IV HASIL DAN PEMBAHASAN

A. Hasil Pengelolaan <i>Dataset</i>.....	41
1. Pre-Processing Data.....	44
2. Penghitungan Frekuensi Kemunculan Kata.....	46

B. Hasil Implementasi Analisis Jaringan <i>Text Co-Occurance Word</i>	47
1. Graph-tool.....	48
2. NetworkX.....	49
3. Networkit	51
4. SNAP	52
5. Igraph.....	54
6. Rustworkx.....	56
C. PENGUJIAN.....	58
1. Pengujian Waktu Eksekusi (<i>Execution Time</i>).....	59
2. Pengujian Penggunaan Memori (<i>Memory Usage</i>).....	60
3. Analisis Data.....	63
a. Nilai Mean (Rata-Rata).....	63
b. Nilai Standard Deviasi	64
c. Nilai <i>Margin of Error</i>	64
d. Boxplot	67
4. Uji Hipotesis	69
a. Uji Normalitas	69
b. <i>Repeated Measures ANOVA</i>	76
c. Uji Friedman.....	72
d. Uji T Berpasangan	72
e. Uji Wilcoxon	75
f. <i>Effect Size</i> dan <i>Satistical Power</i>	75
g. Korelasi Pearson	75
h. Korelasi Spearman	76
5. Analisis Hasil.....	76
BAB V PENUTUP	
A. Kesimpulan	81
B. Saran	81
LAMPIRAN	
Hasil <i>Scrape Datasets</i> Scopus	xxiii
CURRICULUM VITAE	xxix

DAFTAR TABEL

Tabel 1: Sampel Hasil Eksperimen Pengukuran Waktu Eksekusi	35
Tabel 2: Sampel Hasil Eksperimen Pengukuran <i>Current Memory</i>	37
Tabel 3: Sampel Hasil Eksperimen Pengukuran <i>Peak Memory</i>	38
Tabel 4: Spesifikasi Laptop	40
Tabel 5: Hasil Pengujian Waktu Eksekusi dalam Detik.....	59
Tabel 6: Hasil Pengujian <i>Current Memory Usage</i> dalam Megabyte (MB).....	61
Tabel 7: Hasil Pengujian <i>Peak Memory Usage</i> dalam Megabyte (MB)	62
Tabel 8: Rata-Rata Waktu Eksekusi	63
Tabel 9: Rata-Rata Penggunaan Memori.....	64
Tabel 10: Standard Deviasi Waktu Eksekusi	64
Tabel 11: Standar Deviasi Penggunaan Memori	64
Tabel 12: <i>Margin of Error</i> Waktu Eksekusi.....	65
Tabel 13: <i>Margin of Error</i> Penggunaan Memori	65
Tabel 14 : Hasil <i>Mean</i> , <i>Standard Deviation</i> dan <i>Margin of Error</i>	66
Tabel 15: Hasil Uji Normalitas Waktu Eksekusi.....	69
Tabel 16: Hasil Uji Normalitas <i>Current Memory</i> dan <i>Peak Memory</i>	70
Tabel 17: Hasil Normalitas Waktu Eksekusi dan Penggunaan Memori.....	71
Tabel 18: Hasil <i>Repeated Measures ANOVA</i> dan Uji Friedman.....	71
Tabel 18: Hasil Uji T Berpapasan Waktu Eksekusi	72
Tabel 19: Hasil Uji T, Uji Wilcoxon, <i>Effect Size</i> dan <i>Statistical Power</i> Waktu Eksekusi.....	72
Tabel 20: Hasil Uji T, Uji Wilcoxon, <i>Effect Size</i> dan <i>Statistical Power Current Memory</i>	73
Tabel 21: Hasil Uji T, Uji Wilcoxon, <i>Effect Size</i> dan <i>Statistical Power Peak Memory Usage</i>	74
Tabel 22: Hasil Korelasi Pearson	76
Tabel 23: Hasil Korelasi Spearman	76

DAFTAR GAMBAR

Gambar 1: Ilustrasi visual graf.....	11
Gambar 2: Ilustrasi <i>Social Network Analysis</i>	14
Gambar 3: Ilustrasi <i>Text Network Analysis</i> dengan <i>Co-Occurance</i>	16
Gambar 4: Ilustrasi Box-Plot	28
Gambar 5: Halaman Search Scopus	42
Gambar 6: Hasil Search Scopus	43
Gambar 7: Menu Export Data Scopus	43
Gambar 8: Cuplikan Hasil <i>Pre-Processing Data</i>	47
Gambar 9: Hasil Visualisasi grafik NetworkX	53
Gambar 10: Hasil Visualisasi Grafik Networkkit.....	55
Gambar 11: Hasil Visualisasi Grafik SNAP.....	57
Gambar 12: Hasil Visualisasi Grafik Igraph	59
Gambar 13: Hasil Visualisasi Grafik Rustworkx	61
Gambar 14: Grafik Hasil Pengujian Waktu Eksekusi	64
Gambar 15: Grafik Hasil Pengujian Penggunaan Memori	67
Gambar 16: Diagram untuk <i>Execution Time MEAN & Margin of Error</i>	69
Gambar 17: Diagram <i>Current Memory MEAN & Margin of Error</i>	70
Gambar 18: <i>Peak Memory MEAN & Margin of Error</i>	70
Gambar 19: Boxplot Waktu Eksekusi	71
Gambar 20: Boxplot Penggunaan <i>Current Memory</i>	72

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR BAGAN

Bagan 1: Alur Penelitian..... 30



DAFTAR LAMPIRAN

Hasil *Scrape* Datasets xxii



BAB I

PENDAHULUAN

A. LATAR BELAKANG

Penggunaan grafik untuk menganalisis data merupakan cara yang paling populer digunakan dalam berbagai macam domain pengetahuan, seperti penggunaan grafik senyawa kimia untuk penemuan obat (Takigawa and Mamitsuka 2013). Dan penggunaan grafik dalam bidang keilmuan informatika paling umum digunakan pada penelitian terkait *Social Network Analysis* (SNA) untuk mendapatkan pola grafik yang menunjukkan topik yang sedang *hot/trend* pada sebuah jaringan sosial besar (Hu et al. 2013). Penggunaan grafik juga merupakan elemen penting dalam membangun sebuah aplikasi *Natural Language Processing* (NLP), sebagai contoh adalah mesin penerjemah berbasis sintaks, *question-answering* berbasis *knowledge graph-based* dan representasi makna abstrak untuk tugas penalaran *common sense reasoning task* (Pradeep T 2022).

Berbagai macam sistem dan algoritma telah berhasil dimodelkan sebagai suatu jaringan yang kompleks. Karena kumpulan data jaringan besar semakin umum pada era *big data* saat ini. Diperlukan pemrosesan grafik yang efisien, didasarkan pada waktu berjalannya dan penggunaan memori pada kebutuhan perangkat lunak. Namun, Dalam kebanyakan praktik, ditentukan karena fleksibilitas dan kesederhanannya (Staudt, Sazonovs, and Meyerhenke 2014). Memahami kebutuhan perangkat lunak selama proyek pengembangan perangkat lunak merupakan tugas yang cukup sulit. Terlebih jika merupakan analisis menggunakan bahasa alami tertentu yang dihiasi dengan ambiguitas dan ketidaklengkapan.

Oleh karena itu, perlu membentuk suatu kesepakatan bersama antara pemangku kepentingan (*stakeholder*) dan pengembang (*developer*) sistem (Fatwanto 2013). Kesepakatan yang bisa dicapai dalam pengolahan grafik, memungkinkan penggunaan

tools/packages grafik yang sesuai dengan kebutuhan dan kesesuaian perangkat lunak, meskipun *tools* yang digunakan menggunakan bahasa pemrograman yang sama yaitu *python*, dan tujuan yang dicapai sama, beberapa *tools* memiliki perbedaan dalam hal fungsionalitas dan efisiensi.

Manfaat yang begitu banyak dan mencakup ruang lingkup yang besar, membuat grafik lebih dipilih untuk dapat memvisualisasikan sebuah ide maupun penelitian ilmiah. Perkembangan informasi dan teknologi yang semakin maju membuat banyak sekali para pengembang maupun peneliti untuk membuat sebuah alat/pustaka (*tools/packages*) yang dapat mengolah grafik dengan baik. *Tools* dan *Package* untuk pengolahan grafik saat ini sudah tersedia pada berbagai macam *platform* dan banyak yang tersedia secara *online* dan gratis dapat diakses oleh siapapun yang ingin menggunakannya. Beberapa alat grafik (*graph tools/packages*) yang begitu populer yang saat ini digunakan seperti NetworkX, Igraph, Networkkit graph-tool dan rustworkX. Alat tersebut tersedia untuk bahasa python. Kinerja, proses dan antarmuka masing-masing alat grafik dapat dikomparasikan (Pradeep T 2022).

Nikolaive (2022) mengemukakan bahwa NetworkX adalah *library* python untuk analisis jaringan yang cepat, mudah dan umum digunakan pada analisis grafik jaringan yang menyediakan berbagai macam fungsi (Nikolaiev 2022). Sedangkan penelitian lain yang dilakukan oleh Staudt et.al (2014) menyediakan perbandingan untuk networkKit, igraph dan graph-tool untuk algoritma grafik dasar pada grafik berarah (*basic graph alghoritms on directed graphs*). Hasil penelitian menunjukkan bahwa networkkit lebih cepat daripada igraph dan graph-tool pada sebagian besar algoritma yang diuji (Staudt et al. 2014). Berbeda dengan Timothy (2019) yang menguji graph-tool, igraph, networkX, networkkit dan SNAP. Mengungkapkan bahwa pustaka yang tersedia saat ini dapat diskalakan dengan baik dan menangani tugas analisis jaringan yang besar. Fokus penelitian yang dilakukan Timothy adalah komparasi kinerja *loading data, single source shortest path, page rank, k-core*

decomposition dan *strongly connected components*. Networkit dan graph-tool menempati posisi teratas dalam kebanyakan tes/pengujian yang dilakukan.

Pada perangkat yang digunakan seperti laptop, komputer desktop (PC) dan smartphone setiap aplikasi yang berjalan menggunakan memori. Sistem operasi dan aplikasi program yang berjalan juga menggunakan memori, sehingga penggunaan memori merupakan faktor krusial yang harus ditangani seefektif mungkin (Pangera and Ariyus 2017). Selain penggunaan memori, waktu eksekusi (*execution time*) merupakan indikator penting dalam melakukan *benchmarking*

Melihat keragaman, keberhasilan dan kemanfaatan penelitian sebelumnya. Maka penelitian ini mencoba mengembangkan dan menguji dengan pendekatan yang berbeda. Algoritma yang akan dirancang khusus untuk menganalisa grafik berdasarkan judul artikel yang tersedia pada *Scopus*. Studi ini berusaha menyediakan pengukuran kinerja berdasarkan waktu eksekusi dan penggunaan memori pada setiap grafik. Penelitian ini membantu pengembang dan peneliti memilih alat grafik python terbaik untuk kebutuhan mereka. Dengan demikian, penelitian ini memberikan dampak positif pada pengembangan ilmu komputer, informatika dan matematika.

Berangkat dari latar belakang yang telah dijelaskan di atas, penulis tertarik untuk melakukan penelitian dengan menggunakan metode kuantitatif dengan topik “Perbandingan Kinerja (*Execution Time* dan *Memory Usage*) pada *Python Graph Packages* untuk Analisis Jaringan Judul Artikel *Scopus*”.

B. RUMUSAN MASALAH

Berdasarkan pemaparan latar belakang di atas, penulis merumuskan beberapa rumusan masalah dalam penelitian yang sedang penulis lakukan, yaitu di antaranya sebagai berikut:

1. Bagaimana perbandingan waktu eksekusi (*execution time*) pada *python graph packages*?
2. Bagaimana perbandingan penggunaan memori (*memory usage*) pada *python graph packages*?

C. BATASAN MASALAH

Adapun yang menjadi batasan masalah dalam penelitian ini, yaitu di antaranya sebagai berikut:

1. *Graph package* yang akan diuji adalah graph yang populer digunakan dalam penelitian analisis jaringan yakni NetworkX, Networkit, SNAP, Igraph dan graph-tool.
2. Metrik yang akan diuji dan dikomparasikan adalah waktu eksekusi (*execution time*) dan penggunaan memori (*memory usage*).
3. Penulis hanya mengambil data dari 10.000 (sepuluh ribu) judul artikel yang dari *scopus.com* dengan kata kunci “network analysis”.

D. TUJUAN PENELITIAN

Berdasarkan rumusan masalah dan batasan masalah yang telah penulis sampaikan sebelumnya, tujuan dari penelitian yang akan dilakukan ini yaitu:

1. Mengetahui hasil perbandingan kinerja (waktu eksekusi) menggunakan *graph packages* yang telah ditentukan.
2. Mengetahui hasil perbandingan efisiensi (penggunaan memori) menggunakan *graph packages* yang telah ditentukan.

E. MANFAAT PENELITIAN

Berdasarkan tujuan penelitian di atas, manfaat penelitian ini di antaranya sebagai berikut:

1. Dapat menerapkan ilmu-ilmu yang telah diperoleh selama kuliah.
2. Dapat berkontribusi dalam penelitian analisis jaringan teks.
3. Diharapkan dapat memberikan dampak positif pada pengembangan ilmu komputer, informatika dan matematika.
4. Secara praktis diharapkan penelitian ini untuk dapat menjadi bahan rujukan maupun pertimbangan bagi pihak pengembang maupun pengguna dalam menggunakan *python graph packages*.

F. SISTEMATIKA PENELITIAN

Adapun penyusunan laporan penelitian ini dibagi dalam beberapa bagian dengan tujuan memudahkan dalam memahami pembahasannya, yaitu sebagai berikut:

BAB I: PENDAHULUAN

Berisi tentang latar belakang masalah, rumusan masalah, batasan penelitian, tujuan penelitian, manfaat penelitian dan sistematika penelitian.

BAB II: TINJAUAN PUSTAKA DAN LANDASAN TEORI

Menguraikan landasan teori yang mendukung dan sebagai dasar penelitian, yaitu teori-teori yang berhubungan dengan topik yang akan dibahas yakni tentang perbandingan kinerja *python graph packages* seperti networkX, networkit, SNAP, igraph, graph-tool dan rustworkX.

BAB III: METODE PENELITIAN

Menyajikan metodologi penelitian untuk menganalisis kinerja pada *python graph packages*, seperti bagaimana cara mengolah data, mengumpulkan data dan memproses data serta hasil apa saja yang akan didapatkan.

BAB IV: HASIL DAN PEMBAHASAN

Bab ini berisi tentang hasil analisis yang sudah dilakukan dari data yang telah dikumpulkan oleh penulis dari sampel simulasi.

BAB V: PENUTUP

Bab ini berisi tentang kesimpulan dan menjawab rumusan masalah yang telah didefinisikan sebelumnya, serta memberikan saran untuk penelitian selanjutnya.

BAB V

KESIMPULAN DAN SARAN

A. KESIMPULAN

Berdasarkan percobaan terhadap enam paket graph Python, dengan menggunakan dataset sebanyak 10.000 judul artikel yang diambil dari jurnal-jurnal akademis yang terindeks Scopus yang diulang sebanyak 35 kali, dapat disimpulkan bahwa urutan performa berdasarkan waktu eksekusi (dari yang tercepat ke yang paling lambat) adalah: 1. Snap, 2. Graph-tool dan Rustworkx (seri), 3. Networkkit dan Igraph (seri), dan 4. NetworkX (seri). NetworkX.

Sementara itu, urutan efisiensi berdasarkan penggunaan memori (dari yang paling efisien hingga yang paling tidak efisien) adalah: 1. graph-tool 2. snap 3. rustworkx dan igraph (terikat) 4. networkX 5. Networkkit.

Ada korelasi moderat dan positif antara waktu eksekusi dan penggunaan memori. Selain itu, terdapat hubungan yang sangat kuat dan positif antara memori saat ini dan penggunaan memori puncak.

B. SARAN

Hasil dari penelitian ini dapat dijadikan dasar dalam mengambil keputusan untuk menggunakan salah satu paket grafik python dengan mempertimbangkan waktu eksekusi dan penggunaan memori. Namun, penelitian ini juga memiliki kekurangan. Penelitian ini masih menggunakan data yang kecil, maka dari itu diperlukan penelitian lanjutan untuk mendapatkan sampel pengujian dan datasets yang lebih besar.

Disarankan untuk penelitian selanjutnya untuk mengembangkan studi kasus yang jauh lebih kompleks untuk menguji paket grafik python, dalam beberapa kasus hasil mungkin akan berbeda.

Komparasi paket grafik masih terbatas hanya untuk paket grafik dalam pustaka Python, untuk memperluas kajian dan keilmuan maka disarankan untuk penelitian selanjutnya untuk juga menguji berbagai paket grafik dalam bahasa pemrograman yang lain.



DAFTAR PUSTAKA

- Abdullah, Ma'ruf. 2015. *Metode Penelitian Kuantitatif*. Aswaja pressindo.
- Anon. 2016. "Text Networks." Retrieved October 15, 2023 (https://sicss.io/2018/materials/day3-text-analysis/text-networks/rmarkdown/SICSS_Text_Networks.html).
- Anon. 2023. "Snap.Py - SNAP for Python." Retrieved October 16, 2023 (<https://snap.stanford.edu/snappy/index.html>).
- Arboleda, Francisco Javier Moreno, Mateo Rincón Arias, and Jesús Antonio Hernández Riveros. 2023. "Performance of Parallelism in Python and C+." *IAENG International Journal of Computer Science* 50(2).
- Baidoo, Charity Yaa Mansa, Winfred Yaokumah, and Ebenezer Owusu. 2023. "Estimating Overhead Performance of Supervised Machine Learning Algorithms for Intrusion Detection." *International Journal of Information Technologies and Systems Approach (IJITSA)* 16(1):1–19. doi: 10.4018/IJITSA.316889.
- Bates, D., and D. Eddelbuettel. 2013. "Fast and Elegant Numerical Algebra Using the RcppEigen Package." *J. Stat. Softw* 52:1–24.
- Bondy, John Adrian, and Uppaluri Siva Ramachandra Murty. 2008. *Graph Theory*. Springer Publishing Company, Incorporated.
- Bradley, Brooke. 2020. "How to Create Co-Occurrence Networks with the R Packages 'Cooccur' and 'visNetwork.'" Retrieved October 5, 2022 (<https://medium.com/analytics-vidhya/how-to-create-co-occurrence-networks-with-the-r-packages-cooccur-and-visnetwork-f6e1ceb1c523>).
- Callon, Michel, Jean Pierre Courtial, and Francoise Laville. 1991. "Co-Word Analysis as a Tool for Describing the Network of Interactions between Basic and Technological Research: The Case of Polymer Chemistry." *Scientometrics* 22:155–205.
- Dhulipala, Laxman, Jessica Shi, Tom Tseng, Guy E. Blelloch, and Julian Shun. 2020. "The Graph Based Benchmark Suite (GBBS)." *Proceedings of the 3rd Joint International Workshop on Graph Data Management Experiences & Systems (GRADES) and Network Data Analytics (NDA)* Article 11.
- Farida, Nurul. 2020. "Analisis Bibliometrik Berdasarkan Pendekatan Co-Word : Kecenderungan Penelitian Bidang Kearsipan Pada Jurnal Khazanah Dan Journal of Archive and Record Tahun 2016 – 2019." *Khazanah: Jurnal Pengembangan Kearsipan* 13(2):19.
- Fatwanto, Agung. 2013. "Software Requirements Specification Analysis Using Natural Language Processing Technique." Pp. 105–10 in *2013 International Conference on QiR*. Yogyakarta, Indonesia: IEEE.

- Freeman, Linton. 2004. "The Development of Social Network Analysis." *A Study in the Sociology of Science* 1(687):159–67.
- Hagberg, Aric, Pieter Swart, and Daniel S Chult. 2008. *Exploring Network Structure, Dynamics, and Function Using NetworkX*. Los Alamos National Lab.(LANL), Los Alamos, NM (United States).
- Hans, Rizal. n.d. "Tutorial Visualisasi Data Excel Box Plot & Whisker." Retrieved October 18, 2023 (<https://dqlab.id/tutorial-visualisasi-data-excel-box-plot-and-whisker>).
- Harinaldi. 2015. *Prinsip-Prinsip Statistik Untuk Teknik Dan Sains*. Jakarta: Erlangga.
- Hidayat, Anwar. 2012. "Tutorial Uji T Paired dengan SPSS." *Uji Statistik*. Retrieved October 18, 2023 (<https://www.statistikian.com/2012/07/uji-t-paired-dengan-spss.html>).
- Hidayat, Anwar. 2013. "Uji F dan Uji T." *Uji Statistik*. Retrieved October 18, 2023 (<https://www.statistikian.com/2013/01/uji-f-dan-uji-t.html>).
- Hu, Xia, Lei Tang, Jiliang Tang, and Huan Liu. 2013. "Exploiting Social Relations for Sentiment Analysis in Microblogging." Pp. 537–46 in *Proceedings of the sixth ACM international conference on Web search and data mining*.
- J. Wu, V. Srinivasan, and A. Thomo. 2019. "Graph-XLL: A Graph Library for Extra Large Graph Analytics on a Single Machine." Pp. 1–7 in *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)*.
- Krasanakis, Emmanouil, Symeon Papadopoulos, Ioannis Kompatsiaris, and Andreas L. Symeonidis. 2022. "Pygrank: A Python Package for Graph Node Ranking." *SoftwareX* 20:101227. doi: 10.1016/j.softx.2022.101227.
- Lin, Timothy. 2019. "Benchmark of Popular Graph/Network Packages." Retrieved February 5, 2022 (<https://www.timlrx.com/blog/benchmark-of-popular-graph-network-packages>).
- MacFarland, Thomas W., Jan M. Yates, Thomas W. MacFarland, and Jan M. Yates. 2016. "Wilcoxon Matched-Pairs Signed-Ranks Test." *Introduction to Nonparametric Statistics for the Biological Sciences Using R* 133–75.
- Munir, Rinaldi. 2007. *Matematika Diskrit*. Bandung: Penerbit Informatika.
- Nikolaiev, Dmytro. 2022. "Graphs with Python: Overview and Best Libraries." Retrieved February 24, 2023 (<https://towardsdatascience.com/graphs-with-python-overview-and-best-libraries-a92aa485c2f8>).
- Oktaviani, Mitha Arvira, and Hari Basuki Notobroto. 2014. "Perbandingan Tingkat Konsistensi Normalitas Distribusi Metode Kolmogorov-Smirnov, Lilliefors, Shapiro-Wilk, Dan Skewness-Kurtosis." *Jurnal Biometrika Dan Kependudukan* 3(2):127–35.
- Pangera, Abas Ali, and Dony Ariyus. 2017. "Manajemen Memory." 1–6.

- Paranyushkin, Dmitry. 2019. "InfraNodus: Generating Insight Using Text Network Analysis." Pp. 3584–89 in *The World Wide Web Conference*. San Francisco CA USA: ACM.
- Peixoto, Tiago P. 2020. "Graph-Tool Performance Comparison." Retrieved October 2, 2023 (<https://graph-tool.skewed.de/performance>).
- Pradeep T. 2022. "All About Popular Graph Network Tools in Python." Retrieved December 2, 2023 (<https://www.analyticsvidhya.com/blog/2022/04/all-about-popular-graph-network-tools-in-natural-language-processing/>).
- rustworkx Contributors. 2021. "Rustworkx 0.13.2." Retrieved October 16, 2023 (<https://qiskit.org/ecosystem/rustworkx/>).
- Sedighi, Mehri. 2016. "Application of Word Co-Occurrence Analysis Method in Mapping of the Scientific Fields (Case Study: The Field of Informetrics)." *Library Review* 65(1/2):52–64. doi: 10.1108/LR-07-2015-0075.
- Simplilearn. 2023. "Python Timeit()." Retrieved August 3, 2023 ([https://www.simplilearn.com/tutorials/python-tutorial/python-timeit#:~:text=The%20Python%20timeit\(\)%20function%20is%20used%20to%20find%20the,default_timer](https://www.simplilearn.com/tutorials/python-tutorial/python-timeit#:~:text=The%20Python%20timeit()%20function%20is%20used%20to%20find%20the,default_timer).)).
- Spiegel, Murray R., and Larry J. Stephens. 2007. "Statistik Edisi Ketiga." *Jakarta: Erlangga*.
- Staudt, Christian, Aleksejs Sazonovs, and Henning Meyerhenke. 2014. "NetworKit: An Interactive Tool Suite for High-Performance Network Analysis." *Arxiv.Org*.
- Stewart, David B. 2006. "Measuring Execution Time and Real-Time Performance." *Embedded Systems Conference*.
- Sugiyama, Mahito, M. Elisabetta Ghisu, Felipe Llinares-López, and Karsten Borgwardt. 2018. "Graphkernels: R and Python Packages for Graph Comparison." *Bioinformatics* 34(3):530–32. doi: 10.1093/bioinformatics/btx602.
- Sugiyono. 2015. *Metode Penelitian Kuantitatif, Kualitatif Dan R&D*. Bandung: Alfabeta.
- Takigawa, Ichigaku, and Hiroshi Mamitsuka. 2013. "Graph Mining: Procedure, Application to Drug Discovery and Recent Advances." *Drug Discovery Today* 18(1–2):50–57.
- Tsvetovat, Maksim, and Alexander Kouznetsov. 2011. *Social Network Analysis for Startups: Finding Connections on the Social Web*. O'Reilly Media, Inc.
- Zehra, Farzeen, Maha Javed, Darakhshan Khan, and Maria Pasha. 2020. "Comparative Analysis of C++ and Python in Terms of Memory and Time."

Zhang, Zheng, Pierre Zweigenbaum, and Ruiqing Yin. 2018. "Efficient Generation and Processing of Word Co-Occurrence Networks Using Corpus2graph." Pp. 7–11 in *Proceedings of the Twelfth Workshop on Graph-Based Methods for Natural Language Processing (TextGraphs-12)*. New Orleans, Louisiana, USA: Association for Computational Linguistics.

