

Penerapan Algoritma BFS dalam Social Network Analysis pada Github

William - 13518138
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
13518138@std.stei.itb.ac.id

Abstrak - Dengan semakin berkembangnya teknologi – teknologi serta jaringan komunikasi di bidang Ilmu Informatika, semakin banyak hubungan hubungan sosial yang dilakukan di luar hubungan sosial fisik. Salah satu contoh Social Network yang pada era komunikasi modern saat ini adalah melalui Github. Di Github, anda bisa mem-follow para developer lain untuk mengetahui karya – karya apa yang telah ia lakukan, serta developer lain yang memfollow-nya. Analisis ini dapat dilakukan dengan memanfaatkan algoritma Breadth First Search. Dalam makalah ini akan dijelaskan bagaimana algoritma Breadth First Search dapat dimanfaatkan untuk melakukan Social Network Analysis terhadap suatu akun ataupun repo dalam Github.

Kata Kunci—Breadth First Search, Github, Social Network Analysis, follow, repo, akun.

I. PENDAHULUAN

Social Network Analysis merupakan salah satu kegiatan analysis yang sudah mulai populer akhir akhir ini. Berbagai platform sosial media menyimpan berbagai informasi seiring dengan bertambahnya pemakaian sosial media dalam kehidupan sehari – hari. Platform semacam Facebook, Twitter, Instagram, Snapchat memiliki pengguna hingga jutaan pengguna setiap harinya. Sosial media tersebut digunakan baik untuk berhubungan sosial, bermain bersama ataupun memberikan info – info menarik sesama teman.

Jejaring sosial (*Social Network*) merupakan suatu struktur sosial yang terdiri dari node-node yang berinteraksi satu sama lain karena banyak penyebab, seperti kesamaan tingkat sosial, pertemanan di dunia nyata, atau melalui interaksi via dunia digital. Tiap simpul tersebut berinteraksi satu sama lain dan berhubungan pada

tiap – tiap simpulnya, sehingga membentuk suatu jejaring sosial yang besar dari hanya sebuah profil ataupun organisasi.

Salah satu sosial media yang sering dipakai oleh para *developer* di dunia namun tidak menyadari bahwa platform tersebut jugalah sosial media adalah Github. Github merupakan salah satu contoh jejaring sosial yang sangat terkenal terutama di kalangan para *developer*. Data – data terkait jaringan sosial di Github dapat diekstraksi untuk menemukan data – data pertemanan sehingga dapat dijadikan sebagai suatu rangkaian jejaring *follower* dalam bentuk graf yang merepresentasikan bentuk hubungan sosial *developer* satu dengan lainnya pada Github.

Salah satu manfaat mem-*follow developer* lain di Github adalah kita dapat mengetahui apa yang ia lakukan terkait akun Githubnya. Aktivitas seperti meng-commit, push, pull, dan sebagainya yang bersifat publik dapat kita amati. Hal tersebut tentunya dapat kita manfaatkan untuk dijadikan sebagai motivasi dan bantuan jika menemukan project yang bertujuan menyelesaikan masalah terkait. Namun, terkadang kita tidak tahu bagaimana menemukan *developer – developer* yang tepat untuk kita *follow*. Tentunya kita bisa saja langsung mem-*follow* para top *developer* di Github, namun tentunya hal tersebut kuranglah efektif mengingat perbedaan latar belakang dan spesifik masalah yang ingin kita pecahkan. Maka dari itu, *Social Network Analysis* yang dilakukan dalam makalah ini dengan memanfaatkan algoritma *Breadth First Search* (BFS) bertujuan untuk mencari rantai jejaring sosial dari para *developer* yang saling terkait satu sama lain, baik itu melalui keterkaitan *repo* yang diamati (*viewed repo*), *developer* yang di-*follow*, maupun pada *stargazers* akun githubnya.

II. LANDASAN TEORI

A. Algoritma Breadth First Search (BFS)

Breadth First Search Algorithm atau sering disingkat sebagai BFS merupakan suatu satu dari algoritma traversal

graf. Algoritma traversal graf sendiri memiliki makna algoritma pengunjungan setiap simpul dalam graf secara sistematis. Terdapat 2 (dua) jenis algoritma traversal graf, yaitu DFS (Depth First Search) yang mengutamakan traversal secara mendalam pada setiap simpul dalam graf dan BFS (Breadth First Search) yang mengutamakan pencarian secara menyeluruh / melebar pada setiap simpul dalam graf.

Dalam *social network analysis* yang dilakukan pada makalah ini, algoritma BFS - lah yang dipakai sebagai algoritma traversal graf jejaring sosial. Alasannya, karena analisis yang dilakukan lebih bersifat untuk menggali informasi ketimbang untuk memverifikasi suatu informasi. Ketika tujuan dari *social network analysis* adalah untuk mencari informasi, maka algoritma BFS lebih cocok dipakai, karena meskipun lebih kurang efisien dibandingkan dengan algoritma DFS, algoritma BFS bersifat *complete* sehingga kita dapat dengan mudah mencari informasi yang dibutuhkan.

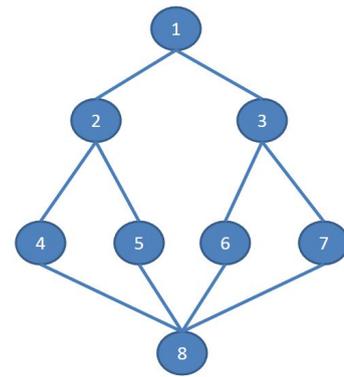
Dalam pencarian solusi dengan algoritma traversal graf sendiri, terdapat 2 (dua) macam pendekatan, yaitu :

1. Graf Statis
Graf yang sudah terbentuk sebelum proses pencarian dilakukan.
2. Graf Dinamis
Graf yang terbentuk saat proses pencarian dilakukan.

Dalam *social network analysis* yang akan dilakukan pada makalah ini, penulis akan menerapkan prinsip graf statis. Hal ini dilakukan dengan memberikan batas pada *depth* pertemanan / *follower* yang akan dianalisis. Dengan memberikan batasan tersebut, graf jejaring sosial dapat diekstraksi terlebih dahulu dengan memanfaatkan scraper akun github, yang nantinya akan di traversal dengan memanfaatkan algoritma traversal graf.

Simpul akar dalam algoritma traversal graf untuk *social network analysis* dalam makalah ini dimulai dari simpul akun Github yang ingin dianalisis (anggap sebagai simpul v). Algoritma traversal graf dalam diterjemahkan sebagai berikut :

1. Kunjungi simpul v
2. Kunjungi semua simpul yang bertetangga dengan simpul v terlebih dahulu
3. Kunjungi simpul yang belum dikunjungi dan bertetangga dengan simpul - simpul yang tadi dikunjungi, demikian seterusnya.



Gambar 2.1 Ilustrasi dari algoritma traversal graf BFS
Sumber: [informatika.stei.itb.ac.id/~rinaldi.munir/Stmik/2019-2020/BFS-dan-DFS-\(2020\).pdf](http://informatika.stei.itb.ac.id/~rinaldi.munir/Stmik/2019-2020/BFS-dan-DFS-(2020).pdf) diakses pada 26 April 2020 pukul 18:25 WIB

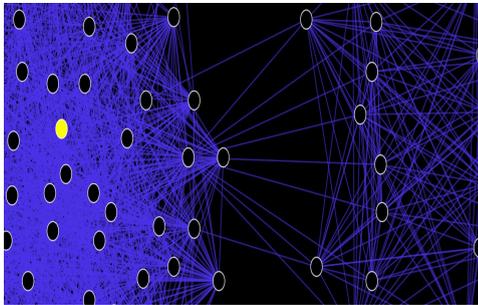
Terdapat beberapa jenis struktur data yang dapat digunakan untuk membantu mempermudah proses traversal graf, sebagai berikut :

1. Matriks ketetanggaan $A = [a_{ij}]$ yang berukuran $n \times n$, $a_{ij} = 1$ jika simpul i dan simpul j bertetangga, $a_{ij} = 0$, jika simpul i dan simpul j tidak bertetangga.
2. Antrian q untuk menyimpan simpul yang telah dikunjungi
3. Tabel Boolean, diberi nama "dikunjungi"
dikunjungi : array[1..n] if boolean
dikunjungi[i] = true jika simpul i sudah dikunjungi
dikunjungi[i] = false jika simpul i belum dikunjungi

B. Social Network Analysis

These visualizations provide a means of qualitatively assessing networks by varying the visual representation of their nodes and edges to reflect attributes of interest
Social Network Analysis (SNA) merupakan suatu proses investigasi struktur sosial melalui penggunaan jejaring / graf dan teori - teori graf. Proses tersebut dicirikan dengan struktur yang berhubungan dan ditandai dengan adanya node (simpul) yang digunakan untuk menyimbolkan suatu peran dalam struktur sosial (aktor individual, orang, ataupun model apapun yang menggambarkan suatu objek dalam struktur sosial) dan edge (sisi) yang digunakan untuk menyimbolkan suatu hubungan antar simpul dalam suatu struktur sosial. Salah satu contoh struktur sosial yang sering digambarkan dalam melalui *social network analysis* adalah struktur sosial dalam media sosial. Contoh - contohnya seperti persebaran meme, sirkulasi informasi, pertemanan ataupun jejaring sosial perbisnisan, jejaring pengetahuan, jejaring sosial. Jejaring - jejaring sosial ini umumnya divisualisasikan melalui sociogram dimana simpulnya digambarkan sebagai titik dan hubungannya yang mengikat dalam jejaring sosial digambarkan dengan

garis yang terhubung. Visualisasi ini memberikan makna kualitatif dalam menghubungkan satu anggota jejaring sosial ke anggota lainnya. Garis (edge) dalam sociograph sendiri merefleksikan atribut yang sedang ingin diketahui dalam social network analysis.



Gambar 2.2 Ilustrasi dari sociograph

Sumber: [https://en.wikipedia.org/wiki/Social_network_analysis#/media/File:Social_Network_Diagram_\(segment\).svg](https://en.wikipedia.org/wiki/Social_network_analysis#/media/File:Social_Network_Diagram_(segment).svg) diakses pada 27 April 2020 pukul 17:42 WIB

III. PENERAPAN SOCIAL NETWORK ANALYSIS DENGAN BFS

Sebelum melakukan social network analysis ini, diperlukan beberapa tools yang akan memudahkan kita. Pertama, penulis menganjurkan untuk melakukan instalasi terhadap Python pada komputer yang akan dipakai.

Selanjutnya, terdapat beberapa modul yang akan diperlukan untuk melakukan analisa social network pada akun github, seperti :

- Instalasi jupyter notebook

```
$ pip install jupyterlab
```

- Instalasi modul PyGithub

```
$ pip install PyGithub
```

- Instalasi modul Networkx

```
$ pip install networkx
```

Dalam melakukan social network analysis pada akun github, pertama kita perlu mendapatkan suatu api key dari github, atau dengan menggunakan username dan password dari suatu akun github.

```
Akses via Username dan Password

from github import Github

USER = '<Insert Username Here>'
PASS = '<Insert the password here>'

client = Github()
```

```
Akses via API Key

from github import Github

# Specify your own access token here

ACCESS_TOKEN = '<insert access token>'

client = Github(ACCESS_TOKEN, per_page=100)
```

Nah, jika setup - setup tersebut telah selesai dilakukan, maka kita dapat melakukan analisis terhadap jejaring sosial akun github maupun repo yang terdapat di github. Namun, sebelum hal tersebut dilakukan, penulis akan mengimplementasikan algoritma BFS kedalam bahasa Python dan membuat beberapa fungsi utilitas lainnya sehingga proses **Social Network Analysis** nantinya dapat dilakukan dengan mudah.

```
BFS Algorithm dalam Bahasa Python

# BFS Algorithm
import collections

# Dasar BFS
def bfs(graph, root):
    print("Starting Graph Traversing with Breadth
    First Search Algorithm : ")
    visited, queue = set(), collections.deque([root])
    visited.add(root)
    while queue:
        vertex = queue.popleft()
        print(vertex)
        for neighbour in graph[vertex]:
            if neighbour not in visited:
                visited.add(neighbour)
                queue.append(neighbour)

# Pemanfaatan algoritma BFS pada social network analysis
stargazers
def bfs_stargazers(graph, root, user_name):
    print("Starting Graph Traversing Repo Stargazers
    with BFS Algorithm : ")
    visited, queue = set(), collections.deque([root])
    visited.add(root)
    while queue:
        vertex = queue.popleft()
        if (vertex == root) :
            show_data_repo(vertex,user_name)
```

```

else :
    show_data_user(vertex)
    for neighbour in graph[vertex]:
        if neighbour not in visited:
            visited.add(neighbour)
            queue.append(neighbour)

# Pemanfaatan algoritma BFS pada social network analysis
Followers
def bfs_follower_network(graph, root):
    """
    Name used is user Username(so it will be unique
    for all user), not Display name
    """
    print("Starting Graph Traversing Follower
    Network with BFS Algorithm : ")
    visited, queue = set(), collections.deque([root])
    visited.add(root)
    while queue:
        vertex, follow = queue.popleft()
        show_data_user(vertex, follow)
        for neighbour in graph[vertex]:
            if neighbour not in visited:
                visited.add(neighbour)
                queue.append((neighbour, vertex))

```

Utility Function untuk Mengkonstruksikan Graf Statis

```

# Utility Function untuk menciptakan struktur graph dalam
# jejaring sosial platform GITHUB
def search_stargazers(repo):
    graph_dict = {}
    graph_dict[repo.name] = [s.login for s in
    repo.get_stargazers()]
    for name in graph_dict[repo.name]:
        graph_dict[name] = [] # Making sure no error
        happens in the program
    return graph_dict

def search_follower(user, depth):
    social_graph = {}
    user_github = client.get_user(user)
    social_graph[user_github.login] = []
    while depth > 0:
        adding_val = []
        for key, value in social_graph.items():
            if len(value) == 0:
                followers_network =

```

```

client.get_user(key).get_followers()
        putback = [follower.login for follower in
        followers_network]
        social_graph[key] = putback.copy()
        adding_val.extend(putback.copy())
        for val in adding_val :
            if (val not in social_graph.keys()):
                social_graph[val] = []
                depth-=1
        return social_graph

```

Fungsi - fungsi tersebut nantinya akan dimanfaatkan saat melakukan **Social Network Analysis**. Untuk analisis jejaring sosial pertama, kita dapat melakukan analisis terhadap jumlah star yang ada pada suatu repo. Dalam hal ini, penulis mengambil contoh repo “**2018-TowerDefence**” yang pernah menjadi Tugas Besar I Mata kuliah Strategi Algoritma.

Analisis Stargazers pada suatu repo

```

# Specify the user and repo of interest
USER = 'EntelectChallenge'
REPO = '2018-TowerDefence'

# Processing the stargazer network by creating a static
# graph of stargazer social network
stargazer_graph = search_stargazers(repo)

# Using BFS Algorithm to show the data of acquired
# information
bfs_stargazers(stargazer_graph, REPO, USER)

```

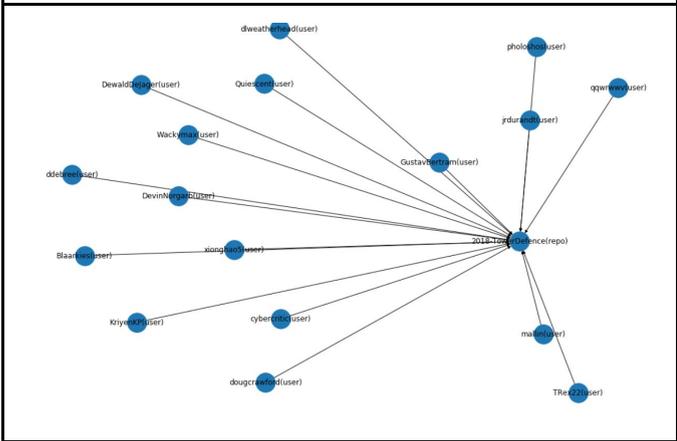
Cuplikan Hasil Analisis :

```

Starting Graph Traversing Repo Stargazers with BFS Algorithm :
Printing Repo Information :
Repo Name : 2018-TowerDefence
Repo Star Count : 17
Repo Language : Java
Repo Topics : []
=====
Printing User Information :
Name : Blaarkies
UserName : Blaarkies
Email : None
Location : South Africa, Pretoria
-----
Printing User Information :
Name : Johannes du Randt
UserName : jdurandt
Email : None
Location : South Africa
-----
Printing User Information :
Name : Jason Chalom
UserName : TRex22
Email : None
Location : Johannesburg, Gauteng, South Africa
-----
Printing User Information :
Name : Mallin Moolman
UserName : mallin
Email : None
Location : None
-----

```

Visualisasi Graf Stargazers Repo :



Nah, tabel di atas memberikan contoh salah satu kegunaan social network analysis, dalam hal ini, kita berhasil menganalisis siapa saja yang memberikan star pada repo Entelect Challenge / 2018-Tower-Defense. Dengan memanfaatkan metode yang sama, kita juga dapat menganalisa follower dari akun user Github tersebut. Pada analisis ini, penulis mengambil contoh yang lebih sederhana, yaitu dengan menggunakan akun github user sendiri sebagai bahan analisis.

```

Analisis Follower pada suatu akun

# Specify a username

USER = 'William9923' # salah satu user yang ingin
di-analisis di github

```

```

DEPTH = 2 # How many of social network depth you want
to traverse

user_graph = search_follower(USER,DEPTH)
bfs_follower_network(user_graph, (USER,None))

```

Cuplikan Analisis Social Network Analysis Follower Github

```

Starting Graph Traversing Follower Network with BFS Algorithm :
Printing User Information :
Name : William Ong
UserName : William9923
Email : None
Location : Bandung
-----
Printing User Information :
Name : Annisa Ayu Pramesti
UserName : nisaprmst
Email : None
Location : Bandung
Follower of : William9923
-----
Printing User Information :
Name : None
UserName : hanifganaitb
Email : None
Location : None
Follower of : nisaprmst
-----
Printing User Information :
Name : William Ong
UserName : William9923
Email : None
Location : Bandung
Follower of : nisaprmst
-----
Printing User Information :
Name : Ananda Yulizar
UserName : anandayulizar
Email : None
Location : None
Follower of : nisaprmst
-----

```

Variable DEPTH yang dimaksud dalam proses analisis user yang diterapkan pada analisis adalah suatu nilai untuk menjangkau seberapa banyak lapisan yang ingin kita analisis. Seperti yang bisa dilihat pada cuplikan hasil analisis di atas, dimana pada layer pertama dapat dilihat bahwa hanya terdapat user “nisaprmst” yang memfollow penulis. Setelah itu, program akan menganalisis setiap orang yang difollow oleh orang pada layer pertama, sebagai layer kedua (DEPTH = 2). Hal ini dapat dilihat dari hasil cuplikan jalannya program, dimana para akun user github yang memfollow akun “nisaprmst” tercetak beserta dengan informasi akunya.

IV. PEMANFAATAN SOCIAL NETWORK ANALYSIS

Dalam kehidupan sehari - hari, social network analysis sendiri sangatlah penting untuk digunakan, untuk hal - hal berikut :

A. Marketing

Social Network Analysis merupakan salah satu metode yang paling mudah dan paling dijamin keaslian datanya untuk mendapatkan informasi -

informasi berharga dalam dunia marketing, seperti untuk mengetahui informasi mengenai suatu komunitas, atau informasi mengenai dampak dari influencer. Segala hal yang menarik di dunia marketing tersebut dapat digunakan dengan baik oleh Community Manager, ataupun Campaign Design untuk menargetkan suatu produk baik barang maupun jasa kepada social network yang tepat.

B. Crime Detection

Social Network selalu membawa informasi yang banyak, baik informasi yang terbuka secara umum maupun informasi yang tertutup. Organisasi - organisasi mungkin dapat menggunakan jejaring sosial untuk mengkomunikasikan sandi - sandi tersebut via jejaring sosial. Social Network Analysis yang lebih menyeluruh atau lebih mendalam (dibanding dengan yang ditawarkan di makala ini) nantinya dapat bermanfaat untuk memecahkan sandi tersebut, sehingga mencegah tindakan kriminal yang mungkin dapat muncul di masa mendatang.

C. Detection of epidemic

Salah satu kegunaan lainnya dari analisis social network adalah untuk mendeteksi adanya penyakit epidemic. Melalui persebaran yang dapat dideteksi dalam analisis jejaring sosial, persebaran penyakit dapat dideteksi sehingga lebih memudahkan pencegahan penyakit tersebut. Salah satu contoh nyata dari hal ini, adalah pemanfaatan social network analysis dalam mencegah pandemic Coronavirus yang sedang terjadi saat ini. Dengan memanfaatkan jejaring sosial, penyebaran Coronavirus juga dapat diketahui dan dianalisis lebih lanjut guna mencegah penyebarannya.

D. Relationship and reputation

Perlu untuk disadari, bahwa social network analysis muncul dari kebutuhan pihak - pihak tertentu untuk mengetahui jenis relasi apa yang dimiliki seseorang dan reputasi apa yang dimilikinya. Kegunaan ini jugalah yang mendasari penulis dalam membuat makalah ini. Dengan memanfaatkan social network analysis, kita dapat mengetahui hubungan sosial apa yang dimiliki seseorang dan seberapa jauh dapat kita mempercayai pihak tersebut (melalui analisis lebih lanjut dengan domain expert). Pada umumnya, hal ini dilakukan di bidang profesional (seperti bagian perbankan).

E. Benchmarking

Merupakan pemanfaatan yang jarang dilakukan yaitu untuk melakukan benchmark terhadap suatu graf, seperti ketika melakukan benchmarking dengan graph RDF management system, ataupun benchmarking social network dalam aplikasi Linked In.

KESIMPULAN

Konsep algoritma traversal graph seperti algoritma Breadth First Search (BFS) sangat banyak dipakai dalam kehidupan sehari - hari. Salah satu kegunaannya adalah untuk melakukan analisis social network yang biasanya berbentuk seperti graph. Melalui algoritma traversal graph ini, kita dapat melihat berbagai hubungan yang terjalin di platform media sosial. Algoritma traversal graph BFS ini juga dapat dimanfaatkan untuk melakukan analisis social network pada platform - platform media sosial lainnya, seperti Facebook, Twitter maupun Instagram. Dengan melakukan analisis terhadap social network, kita dapat mengetahui tingkat sosial seseorang dan dengan siapa orang tersebut berinteraksi. Hal tersebut merupakan suatu data yang amat penting di era digitalisasi saat ini.

VIDEO LINK AT YOUTUBE

Link video : <https://youtu.be/R6hfkQjeXGQ>

PENTUTUP

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Kuasa atas berkat dan rahmat-Nya penulis dapat menyelesaikan makalah ini dengan baik dan pada waktu yang tepat. Penulis mengucapkan terima kasih kepada kedua orang tua serta teman teman yang terus memberikan dukungan baik secara moral maupun doa. Ucapan terima kasih turut penulis sampaikan kepada para dosen yang telah membimbing penulisan makalah ini. Akhir kata, penulis memohon maaf apabila terdapat kekurangan dan kesalahan kata dalam makalah ini. Penulis berharap makalah ini dapat digunakan sebaik - baiknya dan dikembangkan lagi sehingga dapat menghasilkan manfaat untuk masyarakat luas.

DAFTAR PUSTAKA

- [1] Rinaldi Munir, Diktat Kuliah IF2120 : Matematika Diskrit. Bandung: Program Studi Teknik Informatika Sekolah Teknik Elektro dan Informatika Institut Teknologi Bandung, 2006.
- [2] [http://informatika.stei.itb.ac.id/~rinaldi.munir/Stmik/2019-2020/BFS-dan-DFS-\(2020\).pdf](http://informatika.stei.itb.ac.id/~rinaldi.munir/Stmik/2019-2020/BFS-dan-DFS-(2020).pdf). - Materi Algoritma BFS
- [3] <https://www.datacamp.com/community/tutorials/social-network-analysis-python> - Tutorial Social Network Analysis with Python
- [4] https://www.researchgate.net/publication/241060171_Introduction_to_Social_Network_Analysis_Tutorial - Intro to Social Network Analysis
- [5] <https://pygithub.readthedocs.io/en/latest/> - PyGithub Documentation
- [6] <https://networkx.github.io/documentation/stable/> - Networkx Documentation

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 03 Mei 2020

A handwritten signature in black ink, consisting of a large, stylized letter 'W' followed by a smaller, more intricate flourish.

William / 13518138