Influence Scoring

COMP3008

Lauren Ansell

Introduction

Today's topics:

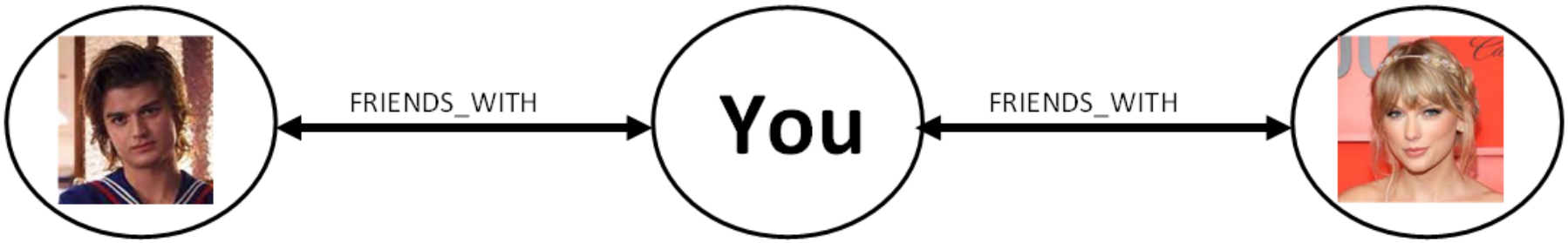
- Influence scoring

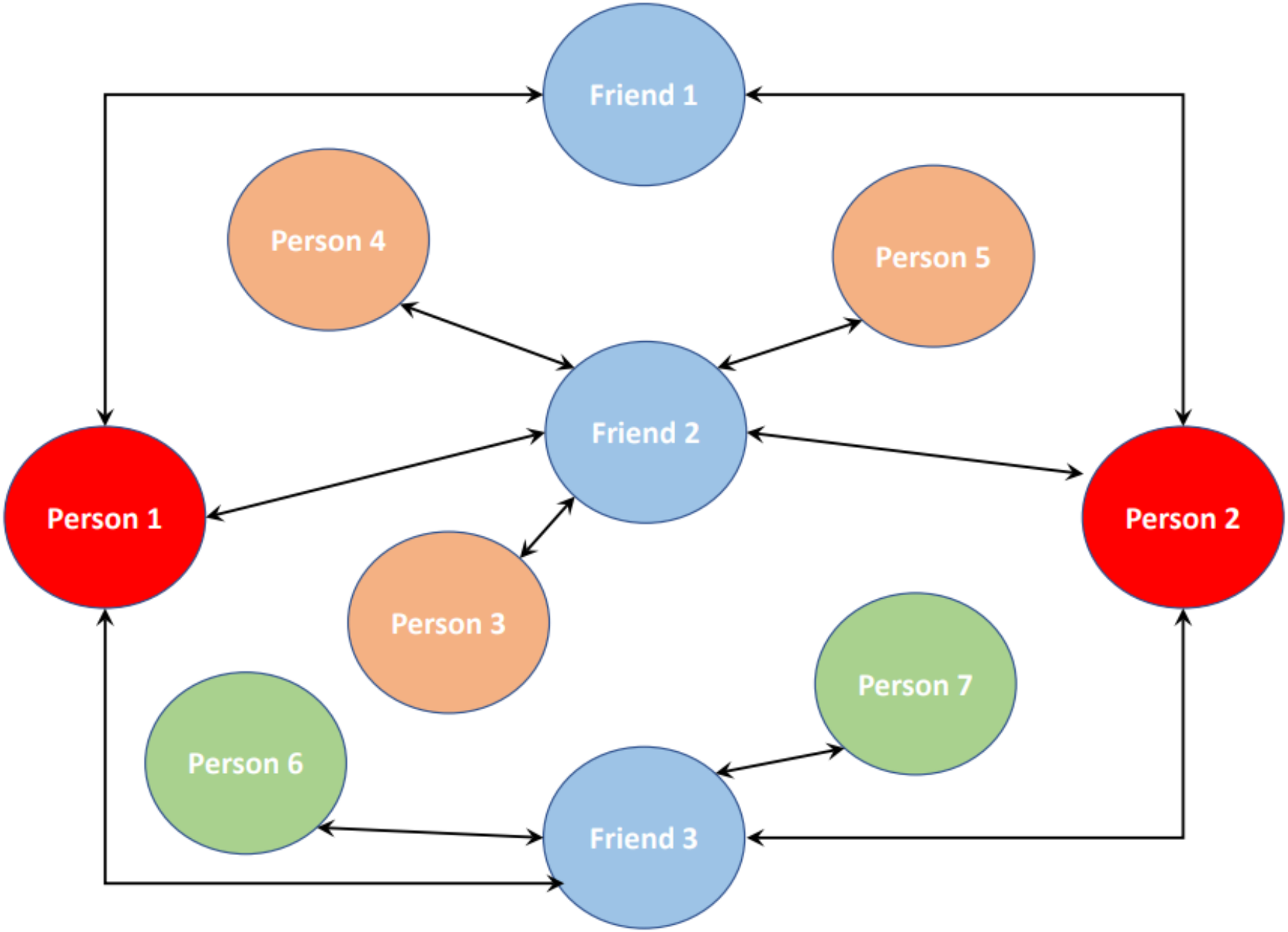
Session learning outcomes - by the end of today's lecture you will be able to:

- Explain and describe what influence scoring is.
- Understand some of the issues affecting influence scoring.

Are All Friends Equally As Important...?

Consider the following hypothetical situation:
You are friends with Steve Harrington and Taylor Swift.
Steve Harrington has only two friends (you and another person).
Taylor Swift has 7,000,000,000 friends.
Steve Harrington and Taylor Swift have no friends in common (apart from you)




Friend 1
Person 4
Person 5
Friend 2
Person 1
Person 2
Person 3
Person 7
Person 6
Friend 3

Influence Scoring

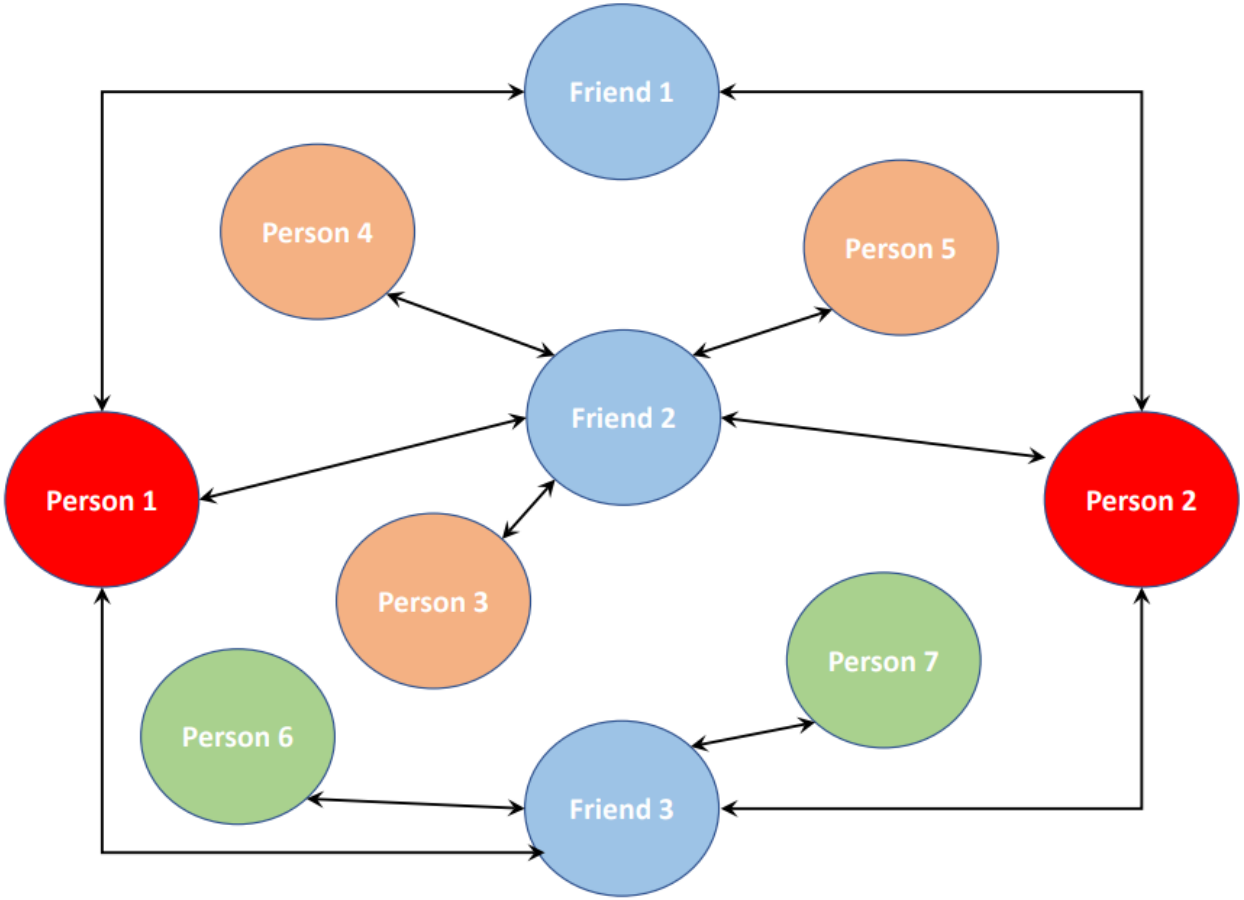
Suppose that person1 and person2 have three friends in common: f_1, f_2, and f_3.

In Collaborative Filtering, the score for person2 as a friend of person1 is $3 = 1 + 1 + 1$.

In Influence Scoring, the score of person2 as a friend of person1 is

$$\frac{1}{\text{numfriends}(f_1)} + \frac{1}{\text{numfriends}(f_2)} + \frac{1}{\text{numfriends}(f_3)},$$

where numfriends(f_n) is the number of friends that f_n has.



How good Person 2 is as a recommendation for Person 1:

$$\frac{1}{\text{numfriends}(f_1)} + \frac{1}{\text{numfriends}(f_2)} + \frac{1}{\text{numfriends}(f_3)}$$

$$\frac{1}{2} + \frac{1}{5} + \frac{1}{4} = 0.5 + 0.2 + 0.25 = 0.95$$

Klout

Klout used social media analytics to rate its users according to online social influence.

It used data from the following pages to create user profiles:

- Bing
- Facebook
- Google+
- Instagram
- LinkedIn
- Twitter
- Wikipedia
- YouTube

Methodology

Klout measured influence by using data points from Twitter.

This information was combined with data from a number of other social network followings and interactions to come up with the Klout Score.

Klout scores were supplemented with three nominally more specific measures, which Klout calls "true reach", "amplification" and "network impact".

Criticism

Several objections to Klout's methodology were raised regarding both the process by which scores were generated, and the overall societal effect.

Klout attempted to address some of these criticisms, and updated their algorithms so that Barack Obama's importance was better reflected.

The site was criticized for violating the privacy of minors, and for exploiting users for its own profit.

Klout was initially criticised for the opacity of their methodology.

In spite of the controversy, some employers made hiring decisions based on Klout scores.

Klout used a hierarchical approach to compute influence scores.

The model also factored in temporal effects and time decay.

Machine learning methods were built using ground truth labels generated for each network.

The overall score for a user is computed by combining the scores from all networks and communities where the user has a presence.

A user was represented as a node in the graph and a set of features was generated.

The features where split into two categories:

1. long-lasting
2. dynamic

The features generated typically have a power law distribution with a long tail.

Non-Negative Least Squares regression was used for model building.

Ranking	ATP	Klout	Forbes	Klout
1	Novak Djokovic	89.54	Hillary Clinton	93.23
2	Roger Federer	90.26	Melinda Gates	83.57
3	Andy Murray	89.50	Mary Barra	77.53
4	Stan Wawrinka	86.86	Christine Lagarde	83.89
5	Kei Nishikori	83.86	Dilma Rousseff	86.84
6	Tomas Berdych	66.69	Sheryl Sandberg	83.18
7	David Ferrer	65.98	Susan Wojcicki	80.04
8	Milos Raonic	82.28	Michelle Obama	87.30
9	Marin Cilic	58.93	Park Geun-hye	81.80
10	Rafael Nadal	82.37	Oprah Winfrey	91.08

Q Score

A measurement of the familiarity and appeal used in the United States.

The Q score was developed in 1963 by Jack Landis.

Q score respondents are given choices for each person or item being surveyed.

$$\text{Positive Q score} : Q_{+} = \frac{\text{favorites}}{\text{known}} \times 100$$

$$\text{Negative Q score} : Q_{-} = \frac{\text{disliked}}{\text{known}} \times 100$$

BitClout

BitClout is an open source blockchain-based social media platform.

Users can post short-form writings and photos and award money to posts they particularly like.

The prices of each account's "creator coin" goes up and down with the popularity of the celebrity behind it

$$\text{price\_in\_bitclout} = 0.003 \times \text{creator\_coins\_in\_circulation}$$

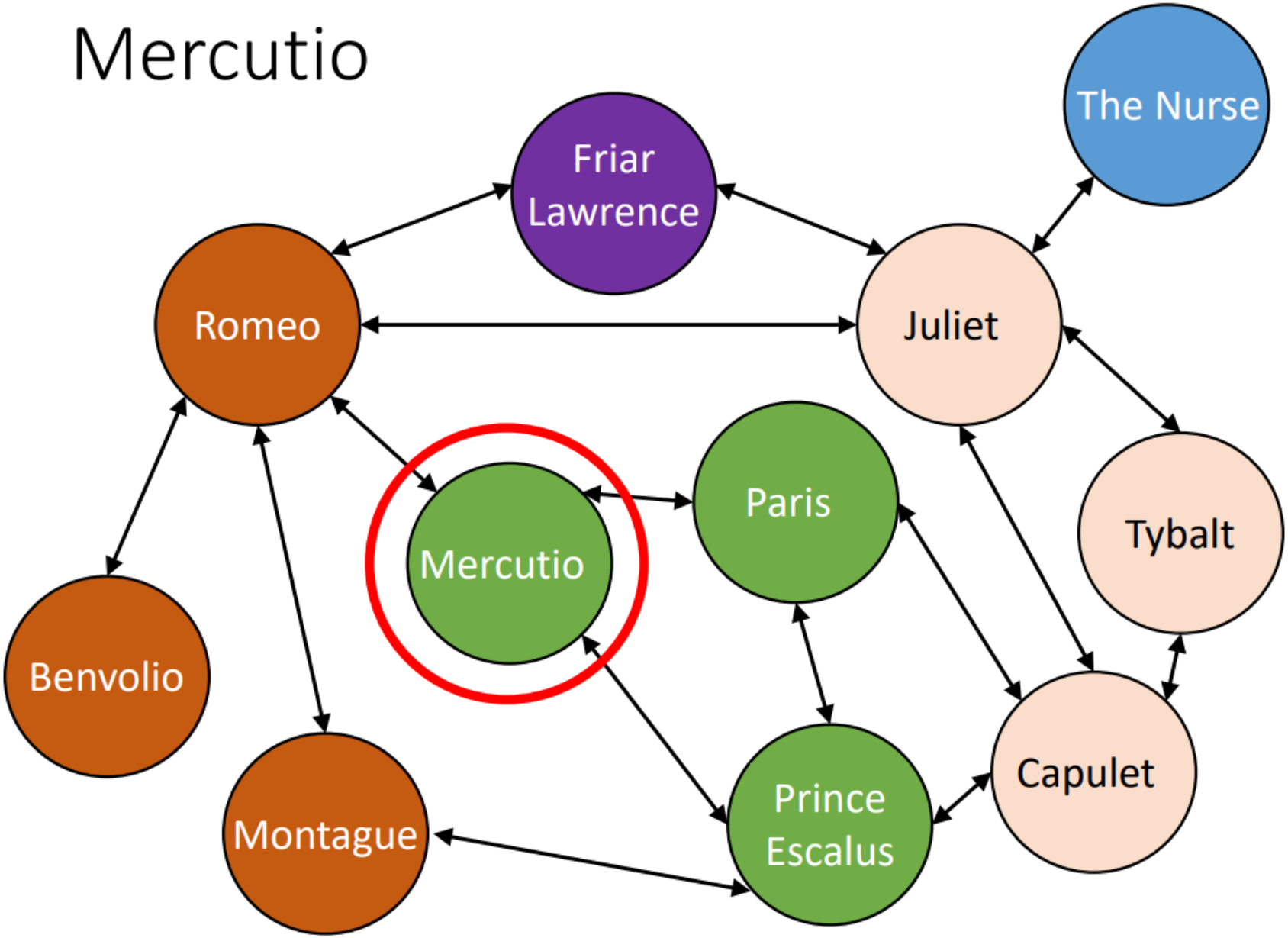
Algorithum Comparison

Does the recommendation algorithm make a difference?

TASK: Using the graph on the next slide, calulate the collaborative filtering score and the influence score between Mercutio and the other characters.

Character A	Character B	C. F.	I. S
Mercutio	Benvolio		
Mercutio	Capulet		
Mercutio	Friar Lawrence		
Mercutio	Juliet		
⋮	⋮	⋮	⋮

Mercutio


The Nurse
Friar Lawrence
Romeo
Juliet
Paris
Tybalt
Mercutio
Benvolio
Capulet
Montague
Prince Escalus

Summary

Influence scoring

- Measures the influence one person has by evaluting their network.
- Based on both the level of engagement and the size of the community.