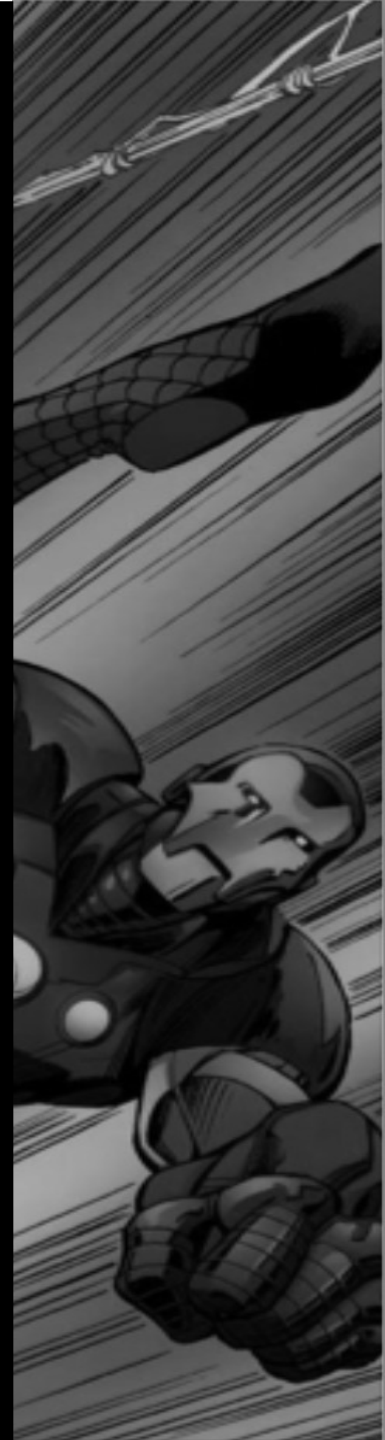# From Comic to Comments:
Marvel Universe Social Network Analysis

MACS 36000

Linhui Wu

Background

Complex network Theory

- Non-trivial topological features

- Community structure, heavy tail

Preferential attachment

- Uneven distributed network

- More connected, more likely to receive connections

Marvel ?

Research Question

Q1: Does characters with higher degrees in Marvel Universe network associate with higher Link prediction accuracy?

Q2: What are the similarities and differences between the communities detected from Marvel Universe network (MU) and the communities detected from comment-based network?

Data

Marvel Universe Network

- Node: character
- Edge: co-appear
- Size: 6246 nodes; 167219 edges
- Heavy tail in degree distribution

Reddit comments

- Subreddit: AskforScience
- Reddit API: Pushshift API
- Top 100 characters with most degrees
- 2000 maximum comments for each node
- Corpus: 64435 pieces of comments
- Vocabulary size: 7850746

Link Prediction

Marvel Universe Network → Adamic-Adar index

↓

Average accuracy for the nodes with same degree ← Exclude the accuracy < 40%

↓

Polynomial Regression → Correlation analysis

↓

Interpretation

Community detection

Eigenvectors + clustering (network-based communities)

- Eigenvectors (all dimensions)
- Elbow method to determine the best k for k-means clustering

LDA Topic modeling (topic-based communities)

- Tokenized and normalized with lucem_illud
- LDA from Genism, coherent score → No.Topics = 10
- Select most frequent topic for each character → edge
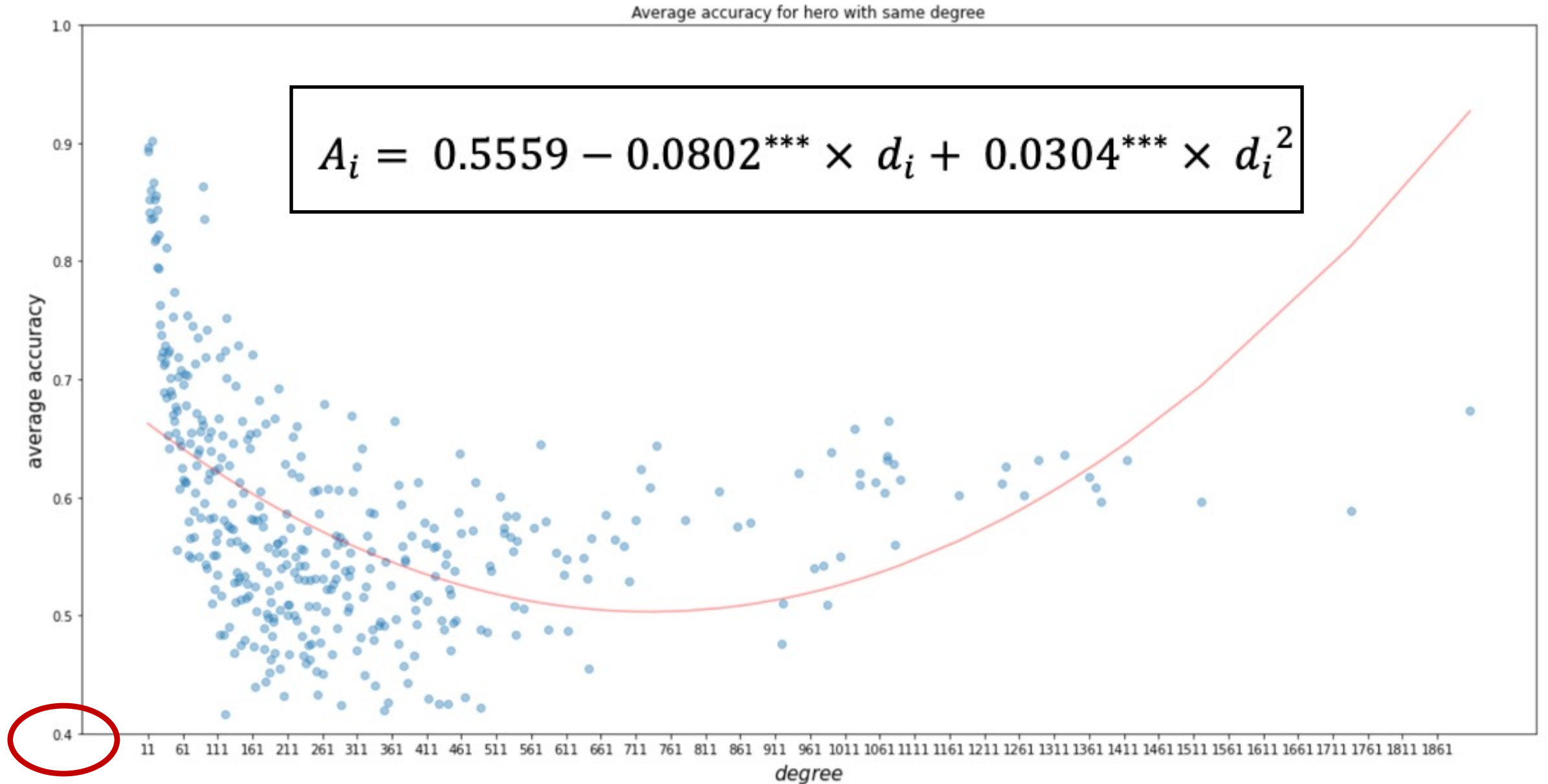
Word embedding (semantic-based communities)

- Pretrained Word2Vec model
- PCA and T-sne for dimension reduction, 300 → 2
- Elbow method to determine the best K=10 for K-means Clustering
- Cosine similarity heatmap

Interpretation of topic

(topic1, "0.010*"marvel" + 0.008*"comic" + 0.008*"year" + 0.008*"earth" + 0.007*"hulk" + 0.007*"time" + 0.006*"aveng" + 0.006*"univers" + 0.006*"charact" + 0.006*"ag") ⟶ **Higher status + political ideology**

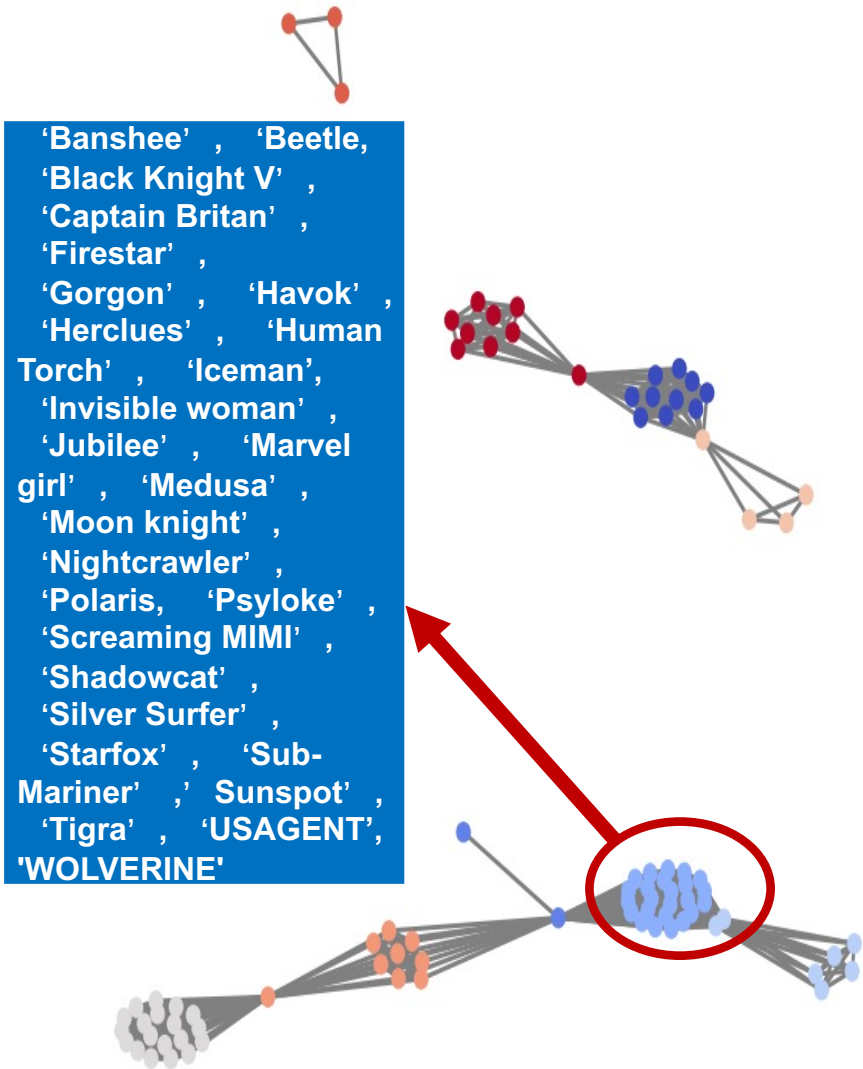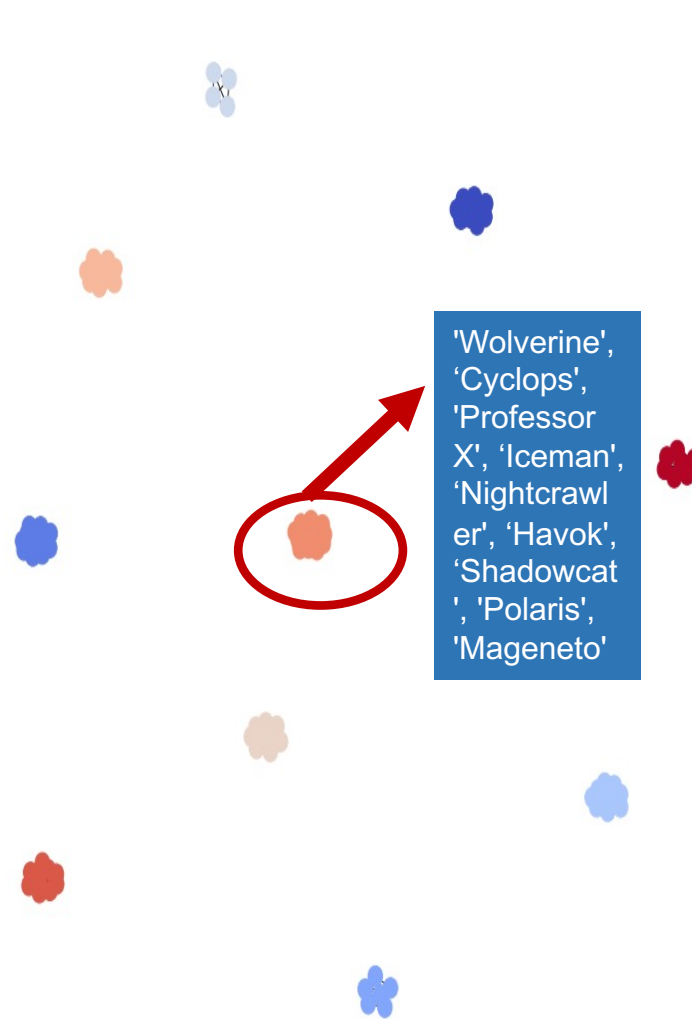| Comment	Probability
Doom is legitimately better than many authoritarians in that his desire to give the people he brutally rules over with a literal and figurative iron fist a better life is sincere. But he will not accept any course of action that involves him doing so by *not* brutally ruling over them with an iron fist and doing as much violence as necessary to establish and enforce that rule.	0.9678
Captain America believes in what America is supposed to represent, not what the country really is. He's perfectly aware of the many fallings of his nation and is ashamed of them. He likes the american dream and the american values Freedom, Security and the pursuit of happiness?	0.9590
He's a veteran. At no point has it shown (that I could remember) that Capt. still has fulfill any military obligations. He's Captain America because he earned that rank and can still use it. If he was promoted post humous, yes, it would still be honored.	0.9499
He genuinely cares about everyone and their rights he just hates Spider-Man being a masked vigilante who answers to no body, he'd very much be called a liberal in real world USA	0.9470

Finding one (Link prediction)

Average accuracy for hero with same degree

$$A_i = 0.5559 - 0.0802^{***} \times d_i + 0.0304^{***} \times d_i^2$$

Finding Two (Community detection)

topic-based communities

'Banshee' , 'Beetle, 'Black Knight V' , 'Captain Britan' , 'Firestar' , 'Gorgon' , 'Havok' , 'Herclues' , 'Human Torch' , 'Iceman', 'Invisible woman' , 'Jubilee' , 'Marvel girl' , 'Medusa' , 'Moon knight' , 'Nightcrawler' , 'Polaris, 'Psyloke' , 'Screaming MIMI' , 'Shadowcat' , 'Silver Surfer' , 'Starfox' , 'Sub-Mariner' ,' Sunspot' , 'Tigra' , 'USAGENT', 'WOLVERINE'

Semantic-based communities

'Wolverine', 'Cyclops', 'Professor X', 'Iceman', 'Nightcrawler', 'Havok', 'Shadowcat', 'Polaris', 'Mageneto'

Network-based communities

'Black Widow', 'Daredevil', 'DR. Doom', 'Firestar', 'Nick Fury', 'Gambit', 'HAVOK', 'WASP'
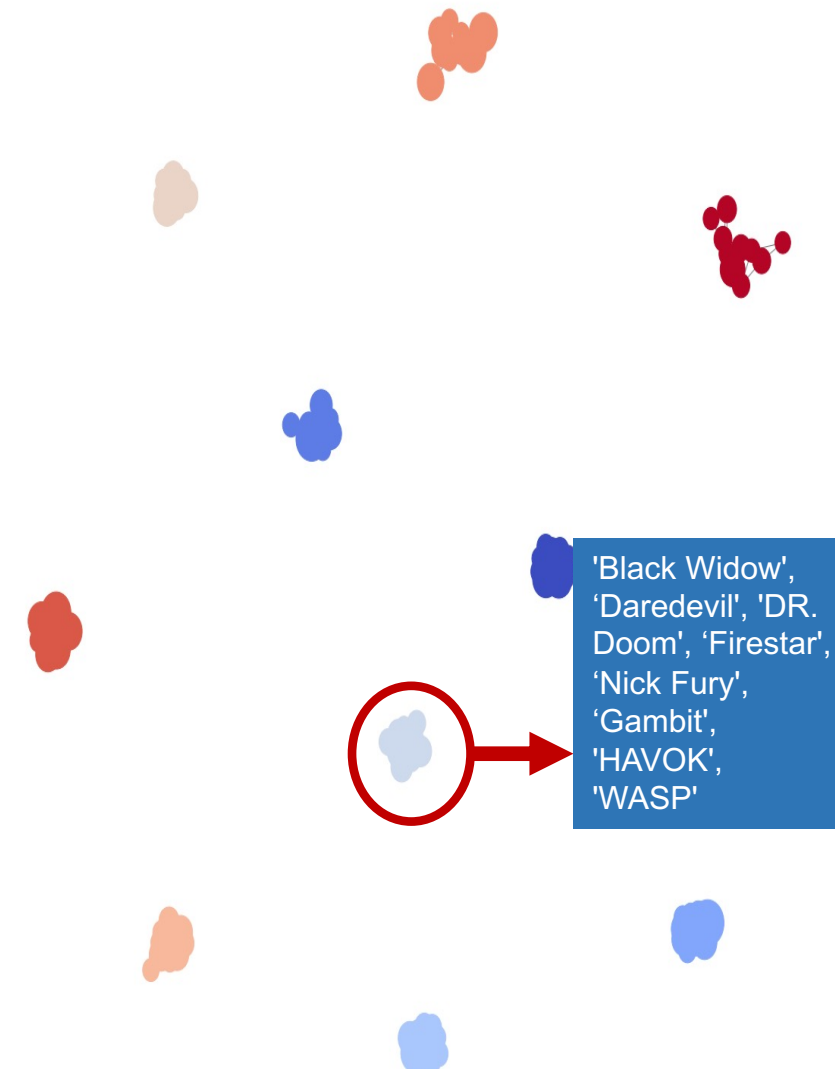
Finding Two (Community detection)
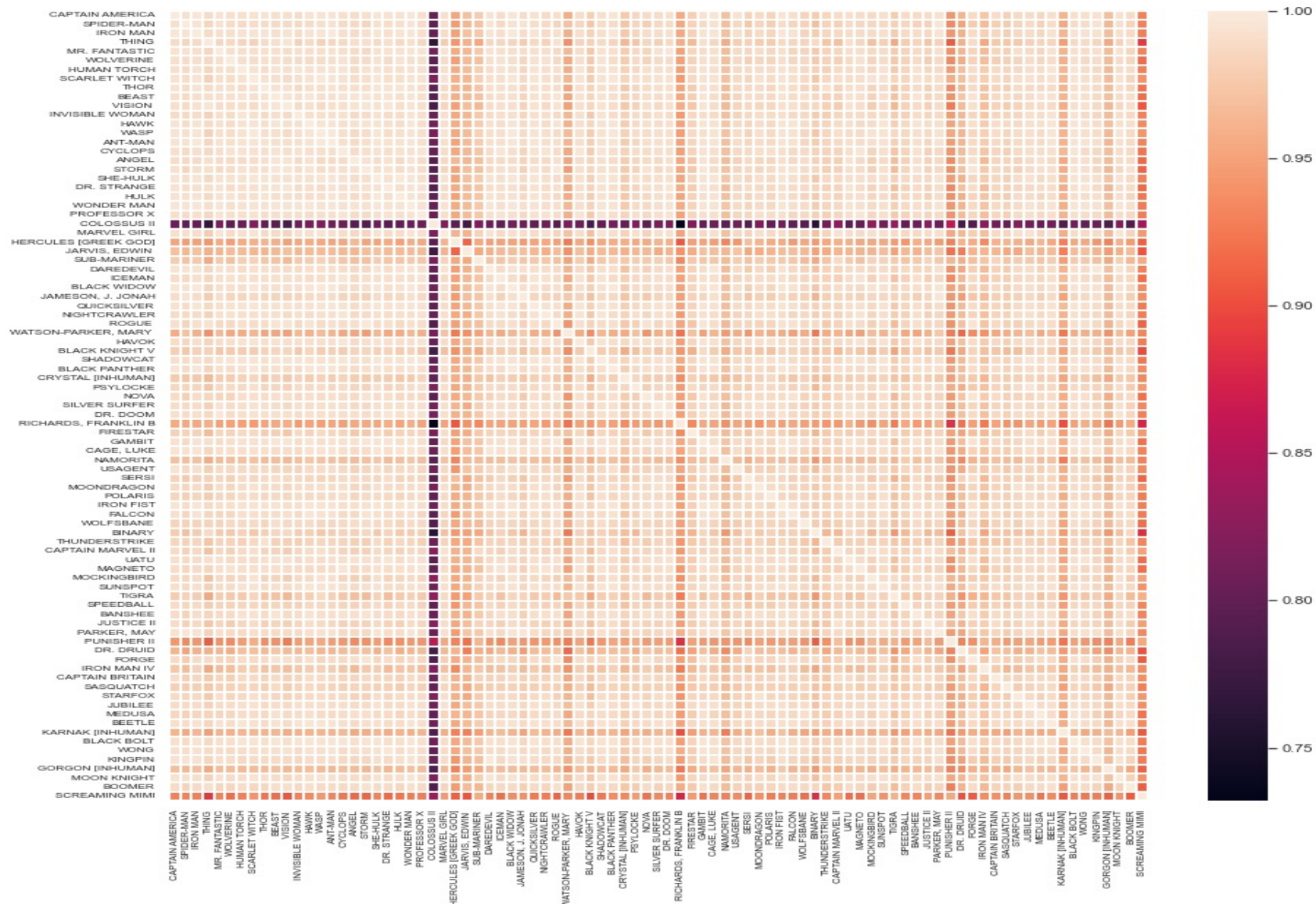
Hero 95% ~ 100%

Villain < 95%

Villain differs from each other

Doctor Druid
Screaming Mimi
Punisher II
……

Summary

1. Link prediction and polynomial regression suggests that future content creation is more likely to focus on popular characters and under-developed characters in Marvel

2. Heroes always represent justice while the characteristic of evil is ever-changing, but evil is an important character for the narrative development; Heroes, Heroic potential, villains

$\downarrow$

Inspiration for future literary creation

Future development

- A muti-facet network considers when, where, who and what
- Simulation of evolution of the network