Network Flow & Dynamic Programming
IE 8753: Fall 2020 Term Project
Zach Douglas Net ID: zd187

1. Introduction

Evaluating node importance is central to a variety of network analysis problems. In terms of social networks, node importance attempts to identify which individuals are more influential to the network as a whole. We are surrounded by many fictitious social networks through entertainment media, a more prominent one being the Marvel Universe (MU). A common area of concern in entertainment media is lack of representation for women and minority groups. While others have studied this lack of representation in entertainment media, there are no known studies which evaluate this representation within a network context. The objective of this project is to identify which characters (nodes) of the MU social network are most significant and relate findings to entertainment media representation.

2. Literature Review

Marvel Comics and the Marvel Cinematic Universe (MCU) are deeply embedded in current American popular culture. Marvel Studios alone has produced 22 films which have grossed $17 billion, making the MCU the largest movie franchise in history [1]. As with any significant entertainment entity, Marvel and its plethora of storylines have the potential to influence communities at scale. However, "with great power comes great responsibility" (Spider-Man).

The effects of comic books are not limited to pure entertainment. Veloso [2] describes their socio-political importance, citing that comic books reinforce ideas that are reflected in current media, thereby playing a role in shaping critical political initiatives (i.e., U.S. participation in WWII, the war on terror). On an individual level, mass media is largely thought to affect how we judge social realities [3]. The implications for how media influences audiences extends beyond entertainment. Myers et al. [4] urge that failing to visually represent marginalized groups in healthcare media prevents those individuals from connecting and instead leaves them feeling invisible.

Mere presence does not constitute sufficient minority representation. Riles et al. [3] investigated character interactions with respect to racial/ethnic identity for four major U.S. television broadcast networks. They [3] found that individuals of color were significantly more likely to interact with White characters and the nature of this interaction was commonly work or crime related rather than romantic or family oriented. Another common criticism is that although minority characters may be more prevalent at present day, their roles are often supporting and are void of the importance given to main characters that are typically White [5].

Previous literature have studied characters in the MU to derive social and cultural insights, but without the utilization of networks [2][6][7][8]. Others have studied the MU as a collaboration network with the focus on comparing and contrasting its features to real-life collaboration networks [9][10]. Gleiser et al. [10] determined that female characters in the MU were not central and did not play a role in connecting communities. At the time of this project, no known study has analyzed the MU network with respect to racial/ethnic representation.

3. Network Description

3.1 Data and Network Development

The network for this study is derived from the Marvel Chronology Project database [11], which details records for which characters appeared in specific comic book issues. This data is originally represented as a bipartite network, with one group of character nodes connected to a different group of comic book nodes. Links between character and comic book represent a character's appearance for that given issue. From this bipartite network, a projected graph is constructed for MU characters which share links if they appeared in a comic book together, thus approximating character relationships.

The original network is a disconnected network consisting of 19 disconnected subnetworks. The 18 smaller subnetworks each consist of 6 nodes or less. The giant component is considered so that methods requiring graph connectedness (e.g., average path length) can be applied. Figure 1 shows the giant component of the network being considered.

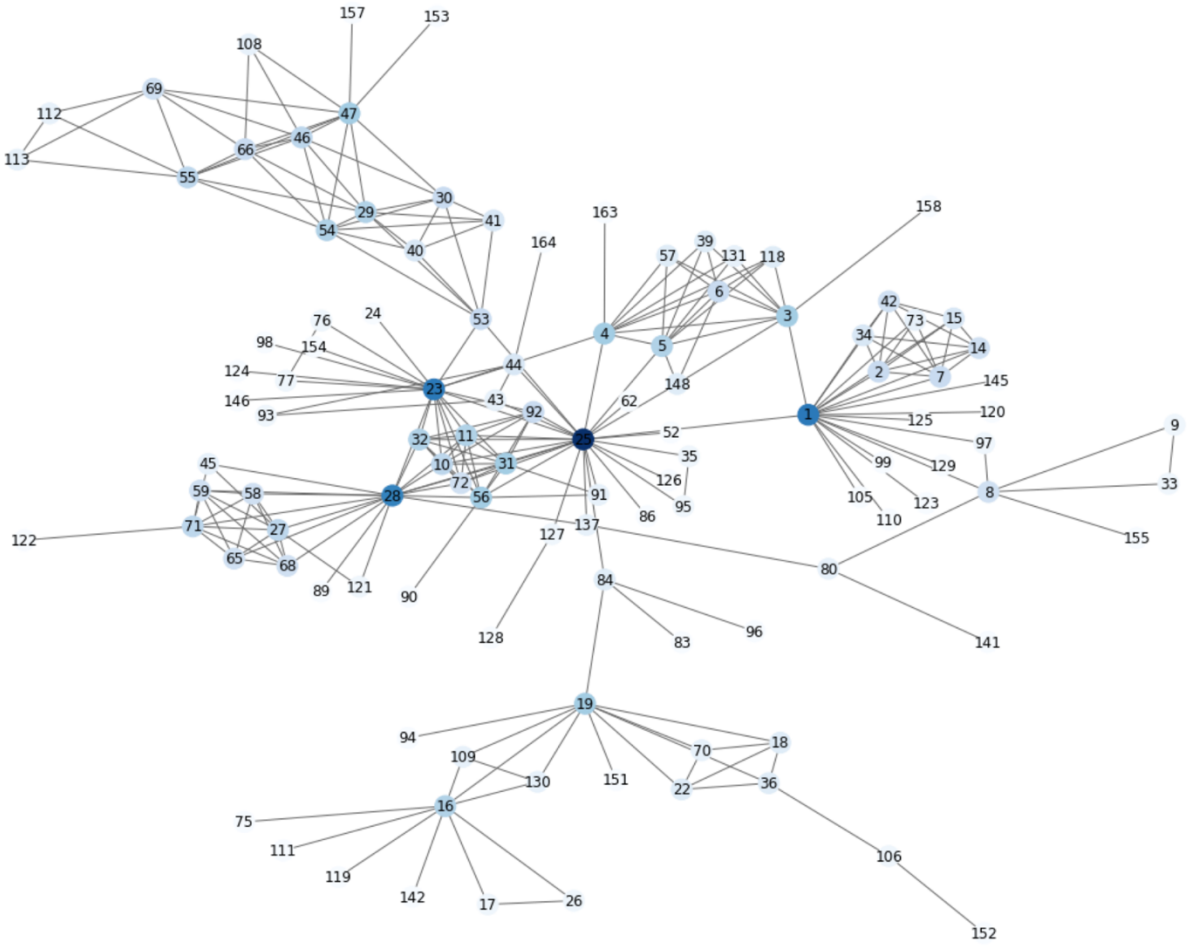

Figure 1: Giant Component Visualization

3.2 Network General Properties

Table 1: Network General Properties

Number of Nodes	113
Number of Edges	249
Density	0.039
Average Shortest Path Length	3.906
Diameter	9
Radius	5
Average Local Clustering Coefficient	0.490
Transitivity	0.442

4. Lessons Learned

4.1 Network Node Centrality

Each algorithm that evaluates network centrality uses different assumptions to rationalize node importance. For the MU network, several algorithms and their respective assumptions apply. Degree centrality signifies character importance since it is feasible that key characters have a relatively high number of relationships compared to less important characters. Closeness centrality is relevant as important characters will be in close proximity to other characters. More important characters require a shorter number of connections to reach other characters.

Betweenness centrality helps to identify characters that serve as important connections for all other characters. Page Rank, Katz, and Eigenvector algorithms work similarly in that they show which characters we are most likely to land upon after *k* random steps [12]. These three algorithms also assume that important characters will be associated with other important characters.

For each of the six node centrality algorithms, a list is created based off of character importance ranking (i.e., a list of the top 10% of characters with the highest degree centrality). The six lists are then combined into a superlist. If a character appears in multiple lists, this repetition is represented in the superlist, where greater repetition corresponds to greater importance. Since each algorithm list is derived independently, the characters between them may differ, resulting in a greater number of characters in the superlist than what is intended to be analyzed (i.e., each algorithm list may identify 10 important characters while the superlist consists of more than 10 characters). To correct for this, characters are sorted by their prevalence in each superlist, and only the top characters of interest are included. Figure 2 shows the top 10%, 15%, 20%, and 25% most important characters using this technique. Character names bolded in red indicate characters who are either female and/or a person of color.

Top 10%	**Top 15%**	**Top 20%**	**Top 25%**
Captain America	Captain America	Captain America	Captain America
Iron Man	Iron Man	Spider-Man	Spider-Man
Thor	Thor	Iron Man	Iron Man
Thing	Thing	Thor	Thor
Spider-Man	Spider-Man	Thing	Human Torch
Hawkeye	**Scarlet Witch**	Hawkeye	Thing
Scarlet Witch	Hawkeye	Mr. Fantastic	Hawkeye
Wasp	Mr. Fantastic	**Scarlet Witch**	Mr. Fantastic
Hulk	**Wasp**	**Wasp**	**Scarlet Witch**
Vision	Vision	Vision	**Wasp**
Ant-Man	Ant-Man	Balder	Vision
Human Torch	Beast	Human Torch	Ant-Man
	Human Torch	Ant-Man	Balder
	Hulk	Odin	Beast
	Wolverine	Beast	Fury
	Dr. Strange	Fury	Cyclops
	Jarvis	Hulk	Professor X
		Wolverine	Odin
		Dr. Strange	Wonder Man
		Cyclops	Hercules
		Professor X	**She-Hulk**
		Nightcrawler	Hulk
		Sub-Mariner	Wolverine
			Dr. Strange
			Storm
			Nightcrawler
			Invisible Woman
			Sub-Mariner

Figure 2: Top 10%, 15%, 20%, 25% Important Marvel Characters (Superlists)

4.2 Network Edge Betweenness Centrality

Table 2: Top 20 Character Relationships

Captain America	Sub-Mariner
Hulk	Sub-Mariner
Captain America	Spider-Man
Captain America	Beast
Captain America	Thor
Dr. Strange	Hulk
Cyclops	Beast
Professor X	Beast
Captain America	Iron Man
Spider-Man	Daredevil
Iron Man	Beast
Hulk	Rick Jones
Captain America	Thing
Jean Grey	Beast
Captain America	Mr. Fantastic
Iron Man	Thor
Spider-Man	Human Torch
Captain America	**Black Panther**
Captain Marvel	Rick Jones
Captain America	Hercules

4.3 Lesson 1: Female and Minority Character Representation

Of the top 10%, 15%, and 20% important characters, all are men with the exception of Scarlet Witch and Wasp. When considering the top 20% important characters, women only make up 8.7% of this group. Among the top 25% important characters, the number of female characters increases from two to five, increasing women representation to 17.9% in this group.

No characters of color appear in any of the top 10%, 15%, or 20% groups. The African American X-Men character, Storm, appears among the top 25% important characters, resulting in a 3.6% representation for people of color in this group.

4.4 Lesson 2: Female and Minority Relationship Representation

Among the top 20 character relationships, two involve Jean Grey and Captain Marvel (White females), and one involves Black Panther (African American male). These relationships represent 15% of the top 20 relationships. Each of the relationships involving females or people of color show these underrepresented characters in relationships with White men.

4.5 Lesson 3: White Male Character Representation

The overwhelming majority of characters across all four superlists are White males. They comprise 83.3%, 88.2%, 91.3%, and 82.1% for each group respectively. Further visualization analysis shows that certain nodes within the network, if removed, would disconnect communities. These nodes are shown in Figure 3, and correspond to characters Spider-Man, Beast, Thor, Hulk, and Daredevil, all of which are White males.

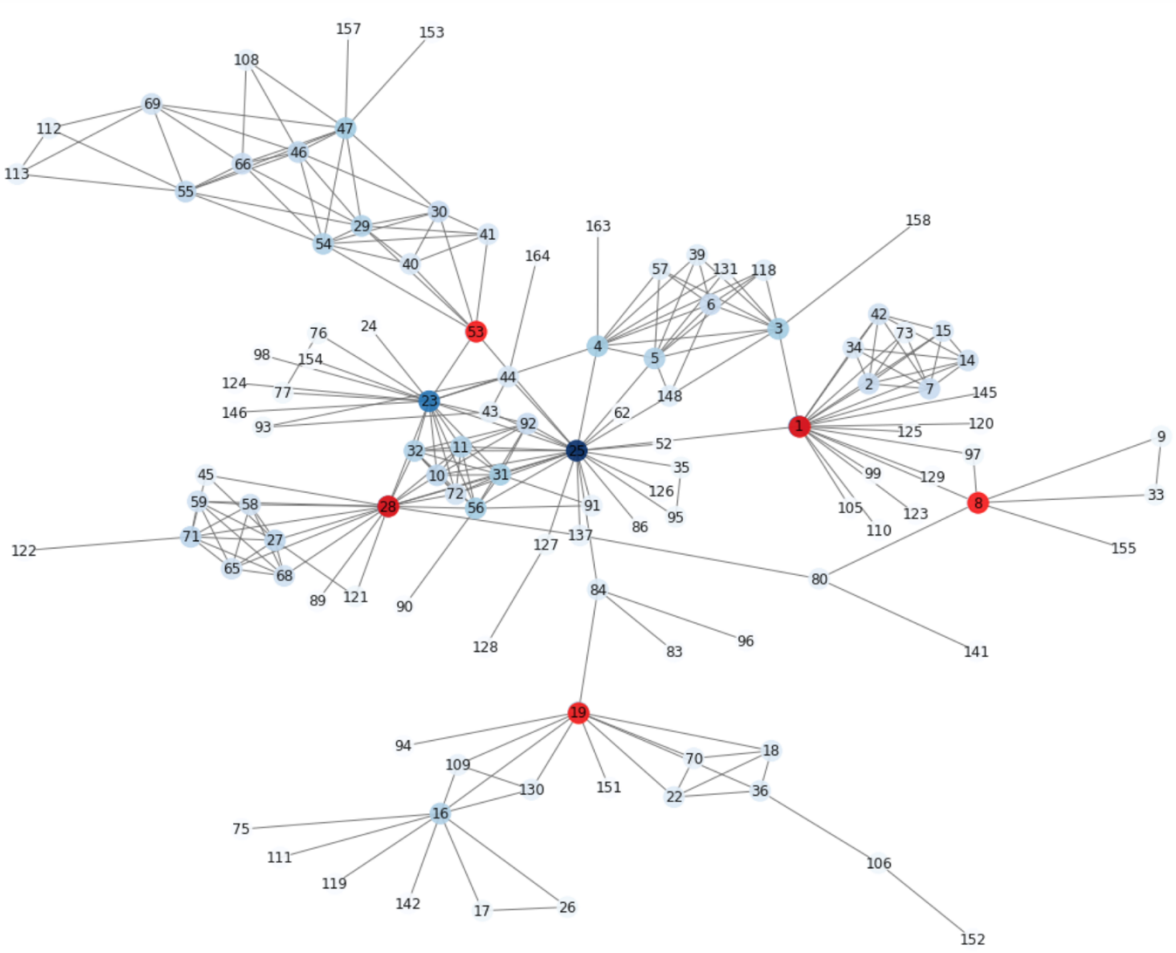

Figure 3: Characters that Disconnect Communities

5. Conclusion and Future Research Directions

While previous studies have examined women and minority group representation in entertainment media, no known study has done so within the context of a network. Network centrality analysis goes beyond determining the presence of underrepresented characters, it allows us to see what role a character plays in a greater community. Additionally, it helps us to quantify character importance. The importance of characters in entertainment media sends a very specific cultural message. The storylines of important characters showcase their relevance. They are often the "glue" that holds everything together. When this relevance is juxtaposed against lesser relevant characters, it sends the message that these less important characters are not as valued. If these unimportant characters are consistently women or people of color, it speaks to a societal paradigm implying that these groups may matter, just not as much as others.

Future considerations and areas for improvement rely heavily on the sophistication of the network being analyzed. Given that the actual MU contains more than 50,000 characters, the MU network in this study is by no means comprehensive [11]. The network could also be improved by including weighted edges, indicating the strength of character relationships. Weights could be derived by accounting for the number of times two characters appear in a comic book together. A multigraph could also be constructed if the nature of character relationships were recognized. The only data attribute for the nodes in this study is the character's name. Character attributes such as gender, religion, race, and age could ease future analyses. Finally, other social networks in entertainment could be analyzed with regard to representation, such as ones seen in TV, movies, books, and DC comics.

References

[1] S. Harrison, A. Carlsen, and M. Škerlavaj, "Marvel's Blockbuster Machine: How the studio balances continuity and renewal," *Harvard Business Review*, pp. 136–145, 2019.

[2] F. Veloso, "Comicbooks as cultural archeology: Gender representation in Captain America during WWII," *Linguist. Hum. Sci.*, vol. 11, no. 2–3, pp. 284–299, 2018, doi: 10.1558/lhs.34737.

[3] J. M. Riles, K. Varava, A. Pilny, and D. Tewksbury, "Representations of Interpersonal Interaction and Race/Ethnicity: An Examination of Prime-Time Network Television Programs," *J. Broadcast. Electron. Media*, vol. 62, no. 2, pp. 302–319, 2018, doi: 10.1080/08838151.2018.1451862.

[4] T. Myers, F. Richardson, and J. E. Chung, "Racial and Ethnic Makeup in Hospital's Social Media and Online Platforms: Visual Representation of Diversity in Images and Videos of Washington, D.C. Hospitals," *J. Health Commun.*, vol. 24, no. 5, pp. 482–491, 2019, doi: 10.1080/10810730.2019.1617807.

[5] M. J. Messineo, "Does advertising on black entertainment television portray more positive gender representations compared to broadcast networks?," *Sex Roles*, vol. 59, no. 9–10, pp. 752–764, 2008, doi: 10.1007/s11199-008-9470-y.

[6] W. Landis, "Ms Marvel, Qahera, and superheroism in the Muslim diaspora," *Continuum (N. Y).*, vol. 33, no. 2, pp. 185–200, 2019, doi: 10.1080/10304312.2019.1569385.

[7] R. A. Griffin and J. P. Rossing, "Black Panther in widescreen: cross-disciplinary perspectives on a pioneering, paradoxical film," *Rev. Commun.*, vol. 20, no. 3, pp. 203–219, 2020, doi: 10.1080/15358593.2020.1780467.

[8] L. do C. Dalbeto and A. P. Oliveira, "Oh My Goddess: Anthropological Thoughts On the Representation of Marvel's Storm and the Legacy of Black Women in Comics," *Comics Grid J. Comics Scholarsh.*, vol. 4, no. 1, 2015, doi: 10.5334/cg.bd.

[9] R. Alberich, J. Miro-Julia, and F. Rossello, "Marvel Universe looks almost like a real social network," no. February 2008, pp. 1–14, 2002, http://arxiv.org/abs/cond-mat/0202174.

[10] P. M. Gleiser, "How to become a superhero," *J. Stat. Mech. Theory Exp.*, no. 9, pp. 1–6, 2007, doi: 10.1088/1742-5468/2007/09/P09020.

[11] R. Chappell, The Marvel Chronology Project, http://www.chronologyproject.com.

[12] M. Marufuzzaman, "Node Importance Part IV (Scaled PageRank)," *Network Flow and Dynamic Programming lecture,* 2020.