‘Marvelualization’ - Tool to Visualize Marvel Comic Characters and their Interactions

Apoorva Ajmani
College of Information Studies
University of Maryland, College Park
aajmani@umd.edu

Ramesh Balasekaran
College of Information Studies
University of Maryland, College Park
rameshb@umd.edu

ABSTRACT

The Comic world is dominated by Marvel Comics with almost 34.20% of market share next only to DC Comics. First published in 1939, Marvel Comics is spread across a variety of genres like crime, horror, mystery, romance, science fiction etc. There are 1400+ characters spread across 77 years of publication. As a result, it is hard to understand the origin, stories, critical events and the relationship between all the characters in the comic universe. With data available from Marvel developer portal and Marvel Wikia our project ‘Marvelualization’ aims to help visualize the marvel comic character’s characteristic and their relationship and interaction.

Keywords: Marvel Comics, Visualization, Comic Characters

1 INTRODUCTION

The competition between two comic book publishers DC Comics almos and Marvel Comics which combinedly holds a market share of almost 70% [1] has given birth to thousands of new characters that attracts wider audience and also targeted demographics. As a result, there is an exponential increase in the quantity of data in the world of comics which originated in the early 1930’s. The rapid increase in quantity of data and it's time-oriented nature adds to the complexity.

Comic lovers now span across multiple generations and face a big challenge in comprehending the origin and history of a character which they start to follow. Typically these fans are not computer experts and require an easy-to-use application that provides narrative visualization by consuming the time-oriented bland categorical data from multiple sources and provides users with direct manipulation that aides in interactive storytelling. Despite interactive visualization being a well-known method to analyze large dynamic datasets, In 2000, Silva, S. F., & Catarci, T observed that enough research had not been conducted on understanding the relationship between the visual techniques and complex time-oriented data [2]. Lu, A., & Shen, H. W. proposed a solution to this problem by developing an interactive storyboard approach that integrates automatic visualization generation methods and interactive adjustment procedures to visualize these time-varying datasets [3].

2 RELATED WORK

2.1 Dynamics of Marvel Comic Universe (MCU) dataset

Marvel Comic Universe (MCU) dataset is complex not only due to its huge volume and time-oriented nature but also due to the relationships between characters across volumes. In 2002, R. Alberich, et.al. visualized the interaction between MCU characters and created a collaboration network and found that the network dynamics resembled that of a real life social network [4].

In 2009, Marvel Chronology Project (MCP) was started to maintain dataset on the appearances of significant MCU characters [18]. Marvel Wikia was created to provide comprehensive information about each character and their interactions. In 2014, Marvel Comics published a developer portal exposing API’s so that general public can create interesting visualizations. In 2014, C. Harris used this API to collect details about the Marvel characters and created visualizations in Tableau [5]. The infographics created from the visualization revealed interesting facts about the Marvel comic world. Even though the visualization provided interesting facts about the dynamics of Marvel characters it still fell short in explaining the story behind.

2.2 Dynamic Network Timeline Visualization

Different visualization techniques like scatter plot, venn diagram, network diagram etc. have been adopted to visualize the MCU dataset. Huge scope for new and innovative visualization still exists due to the dynamic and cumulative nature of the data and the need for presenting the story along with insights. Tim Leong in his book ‘**Super Graphic: A Visual Guide to the Comic Book Universe**’ has explored multiple ways to visualize comic characters such that it not only conveys a narrative story but also looks authentic to the comic lovers [6].

In order to cater to the need to represent MCU’s time-oriented data, dynamic network timeline visualization seems like an optimal approach. The dynamic network incorporates key interactions of characters, their appearances and disappearances over a successive period of time/comic/series. M. C. Ardanuy and C. Sporleder used a similar approach in the field of literary analysis to examine the extent to which a character structure of novel like ‘*David Copperfield*’, ‘*Around the world in 80 days*’, ‘*Peter Pan*’ is indicative of its genre [7]. Dynamic network best represents temporal and spatial data and helps in identifying the concept of ‘*chronotope’* - narrative time-space unit [8].

With the advent of digital comics, interactive visualization in comics has been explored and studied to enhance narrative storytelling. In 2012, Andrews, et al. integrated direct manipulation and branching comic theory to provide a compelling reason for reading digital comics [9]. The same approach can be followed in MCU visualization, by providing an interface to the users to manipulate the narrative and explore the complex and dynamic MCU space. The dominance of the presence of interactive visualization can also be attributed to game mechanics. In 2011, Nicholas Diakopoulos, et al. weighed the benefits of converting the complex dynamic data into playable data that provides an alternative method to structure the story which is not bound by linear arrangement but still provides the benefits of a latent structure [10].

2.3 Social Network Visualization

The main essence of the MCU is the interaction between characters. Crossovers between comic series have captivated the audience since a long period of time and as a result visualizing these interactions between characters would not only help comic lovers to explore stories between the comic characters but also would help in creating a mental map on how these characters are related to each other. Experts in the field of visualization have frequently created 'social network' to visualize the interaction between characters in novels, movies, and comics. In 2012, Apoorv Agarwal et al. used social network graph for analyzing the interaction between characters of 'Alice in Wonderland' and concluded that social networks based on social events provide interesting insights into the role of each character [11].

Traditionally, researchers have used node-link diagrams, arc diagrams and matrix views for visualizing social networks. In 2007, N. Henry et al. observed that 54 out of 55 social network visualization software used node-link diagrams to solve this particular problem but at the cost of visual hierarchy [12]. Node-link diagram also has its own shortcomings in terms of wasted space when applied to 1400+ characters in MCU. J. Heer et al. in their research recommended force directed node-link diagram to be an efficient method since the space required for each node is computed before the visualization is rendered [13]. In 2015, S. Silvia et al. further enhanced the force directed node-link diagram by introducing interaction that applies force iteratively and continuously while the users navigate through the visualization [14]. This type of visualization facilitates flexible design and multiple view visualization that includes storyline view.

J. Heer and B. Shneiderman in 2012 created a taxonomy of tools that support the fluent and flexible use of visualization and expressed sorting or ordering as a fundamental operation for a fluent visualization [13]. Network visualizations do not present straightforward opportunities for sorting. M. Bostock in 2012 used an adjacency matrix with various options to sort the data like cluster, name, frequency etc which resulted in an interesting visualization of interactions between characters in Les Misérables [15]. Sorting by different attributes of data provides various patterns that help in understanding the underlying story. Adjacency matrix also overcomes the problem of giant hairballs of line crossings in the case of large node-link diagrams.

After weighing the benefits and drawbacks of the two different visualizations - Storyline Visualization with Force Directed Layout and Adjacency Matrix, we feel that both would provide complementary insights into the MCU.

2.4 Color and Character Identity

A sequential narrative works best when words, colors, and interaction go hand in hand. Color being a fundamental visual feature can be mapped to various data attributes to reveal interesting insights in a visualization. Talon, Durwin S (2007) explored how emotions, mood, feelings and character identity can be altered by effective color theory [16]. Seeum Lim & Seodaemun-gu's [17] study on Marvel comic characters has also revealed the importance of the use of dominant color in superhero character identity. **In 2015**, before the release of '*Avengers: Age of Ultron*' movie, **Randy** used the Marvel API and observed the color theory of Marvel Comics by visualizing cover pages of all Avengers comics that spanned 50 years. The visualization revealed a gradual shift in color identities to darker colors with the advent of digital comics and movie influence. Following the visualization, The Wall Street Journal published a story titled '***50 Shades of 'Avengers': The Colorful History of Earth's Mightiest Heroes*** summarizing the color dynamics of Avengers. With the recent release of '*Captain America: Civil War*' a similar visualization to understand the change in color dynamics and character identity of Captain America and Iron Man would be interesting.

2.5 Visual Information Analysis

Interpretation of a visualization is as critical as deciding the correct approach for its representation. While developing an interactive visualization for a complex time-oriented data set it is important to understand how the users approach the tool and interpret the rendered visualization. **In 2008, P. Isenberg, et al**. analyzed and presented a framework for interpretation of visualization for non digital setting. A similar framework needs to be studied and developed for interactive digital visualization. For our project we will use a condensed version of the framework involving five processes namely browse, clarify, select, operate and validate.

3 IMPLEMENTATION

3.1 Data Collection

Marvel provides API to share information on Marvel characters, comics and events information. This information was initially collected programmatically using code written in node.js and a local database with the information was formed using excel sheet. The API provided information on each Marvel character, the comics list in which the characters appeared and the details of each comics including cost of comic, creator information, etc. For each comic character Marvel API also provides links to pages in the Marvel Wikia that contains the detailed information of the character. Addition details on the comic characters was web scrapped from the Marvel Wikia website using these links. Additionally, interaction between each characters was computed by creating an edgelist with weight as the number of comics each character appeared in together.

3.2 Initial Sketches

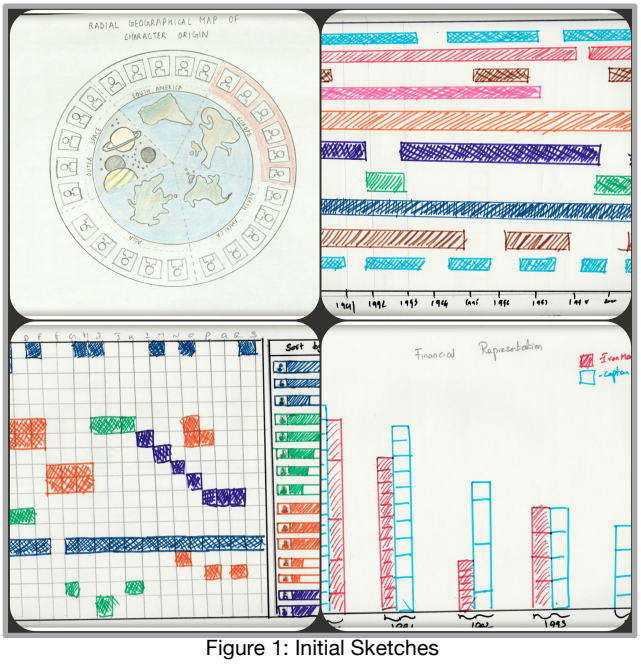

Figure 1: Initial Sketches

Our initial thought process was to create a visualization that would help users to interact and explore the data from Marvel API to understand the MCU. Our aim was to create a narrative visualization that conveys the character's origin, lifespan, popularity, interaction with other characters and and its influence on Marvel's financial success. We explored different types of visualization from scatterplot, geographical, narrative charts to radial and adjacency matrix. With further brainstorming and research on how past researches have used different type of visualization to tackle similar problems, we incorporated the benefits of different type of visualization with multiple interactive visualization elements and accordingly modified our visualizations. Understanding the needs of our visualizations we decided to use d3.js for all of them.

3.3 Geographical Map of Character Origin

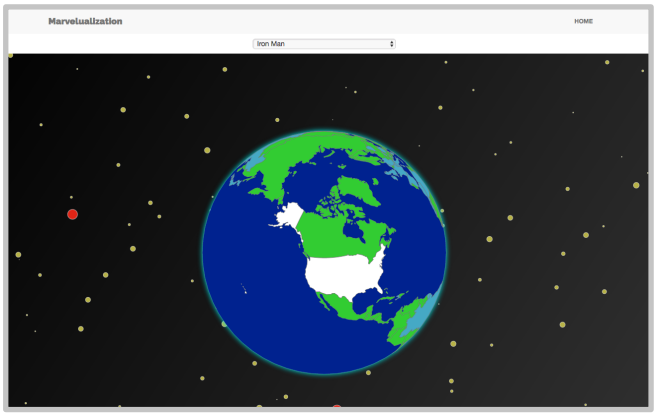

Figure 2: Comic character - USA, Earth

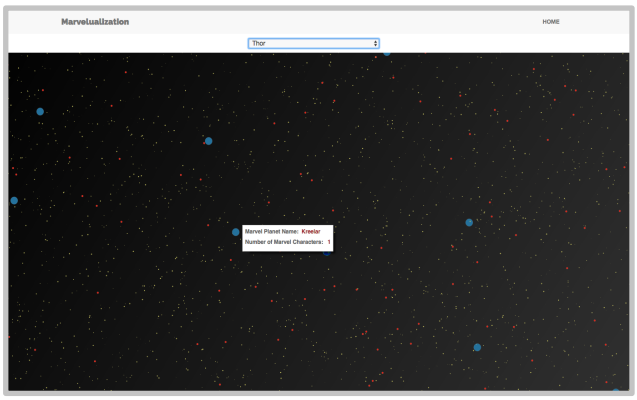

Figure 3: Comic character origin - Outer Space

From the collected data we extracted character's geographical information and created a JSON. The JSON consists of each characters with city of origin, country, location type and character id. In order to visualize the character's origin, d3.js visualization that projects each country's path orthographically and stars in equilateral projection was used. The countries and planets that had at least one Marvel character originated from them was highlighted. This helps in visualizing the countries from which Marvel character's originated. Additionally a dropdown with list of characters in the visualization helps to find the country of origin for each character by selecting the country and rotating the globe to the country. When a character from outer space is selected the map zooms to reveal the entire universe.

3.4 Marvel Character Representation over Timeline

From the collected data, timeline for top 50 Marvel characters was computed and the resulting data was converted into JSON. In order to visualize the introduction, appearance and disappearance of these Marvel characters d3.js timeline was used. Year was plotted on X-axis and Name of Character was plotted on the Y-axis. Figure 4 represents the timeline from 1939-2016. Gaps in each row symbolize the absence of the character during that particular timeframe. This visualization helps to identify and analyze patterns for the lifeline of a character as well as a group in the Marvel Universe. In order to filter the available timeline in regards with the start date and the end date a slider has been placed at the top of the visualization. Upon selection of a particular year the visualization is reloaded with values corresponding to the updated time range. Additionally, on the click of each bar the corresponding start and end date for the particular time span for that character will be displayed.

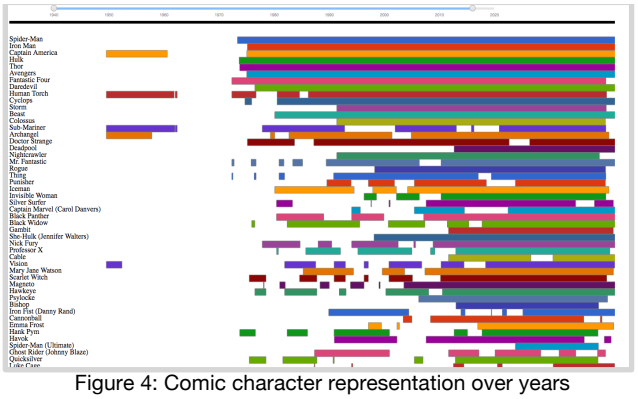

Figure 4: Comic character representation over years

3.5 Interaction Matrix

For the top 50 Marvel characters, edge list between all the characters was created, For each edge, the weight was computed based on the count of comics the two characters appeared together. The resulted dataset was converted into JSON. The JSON was then used to create a adjacency matrix using d3.js. Additionally the adjacency matrix was modified to include a list of characters in the right section. In traditional adjacency matrix the cells are rearranged either by name or group or by attributes of the characters. With the new section created with list of characters users can alter the matrix manually by dragging and dropping the characters in the list. Each section also marks the group to which the character belong and how many comics in total the character has appeared.

The color intensity of each cell denotes the strength of interaction between two characters. Characters which appeared in more comics together had high color intensity. This helps to identify which character plays a major role in interacting with other character or a group. Each cell can be color coded based on group affiliation or dominance. If color by affiliation, cells which both the character belong to same group will be color coded with group's color definition. If color by dominance, cells will be color coded based on dominant character between the two.

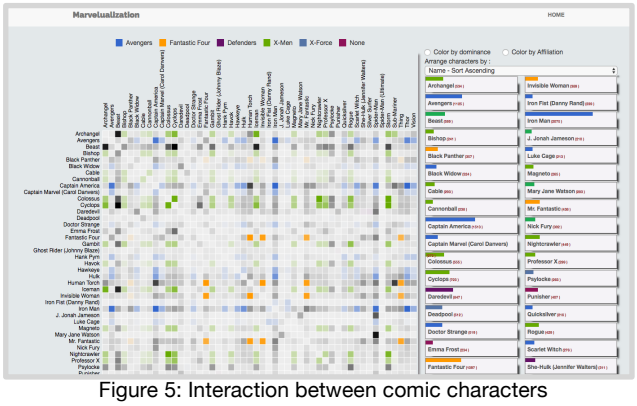

Figure 5: Interaction between comic characters

3.6 Cost Analysis between two Marvel characters

Figure 6: Cost analysis between comic book price of Iron Man and Captain America

From the data collected for Interaction Matrix, the cost for each comic for all Marvel characters was extracted and converted into JSON. In order to represent the price difference as well as the face value for a character over a period of time d3.js stacked grouped bar chart visualization was used. Year was represented on X-axis and Total Cost in USD was represented on Y-axis. Figure 6 represents the cost analysis between Iron Man (Red) and Captain America (Blue). The height of each bar conveys the cost of print price of a comic book for that character. Color intensity signifies a steeper cost difference between the print price of the comic book for that particular year. In order to filter values corresponding to a particular start and end date a slider has been placed at the top of the visualization. On hovering over a bar Character Name, Comic Name and Comic Cost is displayed.

4 RESULTS

4.1 Analysis of geographical map depicting character origin

We observed that Marvel characters were spread across the globe (Americas, Asia and Oceania) and Marvel universe. USA lead the list with 405 Marvel characters followed by UK, Russia, Japan and Asgard.

4.2 Analysis of Marvel Character Representation over Timeline

During the period of 1960 - 1970 a sharp rise in the number of new Marvel characters is witnessed. From the day of first appearance to present only Spider Man, Iron Man, Hulk and Thor had made their appearance consistently. Amongst the top 50 characters, Spiderman Ultimate and Deadpool were the most recent additions to the Marvel Universe with steady increase in the comics from first year of appearance to present. Captain America made his first appearance in 1941, twenty years before any other Avenger character made his appearance thereby becoming the founder of the group. Appearance of the original characters of 'Fantastic Four' had been gradually reducing with Mr. Fantastic last appearing in 2014, Human Torch, Thing and Invisible Woman in 2012. It is interesting to find the character Vision who appeared in 1940 made a reappearance only after approximately twenty five years.

4.3 Analysis of Interaction Matrix

It was observed that each rearrangement resulted in interesting patterns that provided new information about the comic characters. When the matrix is sorted by groups, X-Men (green), Avengers (blue) and Fantastic Four (orange) appeared to be the strongest groups consisting of the majority of the top 50 Marvel characters. 'Avengers' turned out to be biggest group with 15 Marvel characters. In 'Avengers', Iron Man had strong connections with Captain America, Thor and Spiderman. Characters affiliated with 'Fantastic Four' displayed similar level of interaction with each other except Black Panther who had more interactions with people outside the group. This can be justified by the transition of Black Panther form 'Fantastic Four' to 'Avengers'.

When the matrix is sorted by frequency, 'Avengers' led the list followed by 'Fantastic Four' and 'X-Men'. In each group we observed that some characters were very prominent for example 'X-Men', Cyclops, Storm and Colossus have high interaction strength with other 'X-Men'. Apart from 'Avengers', Spider Man had highest interaction with Mary Jane Watson explained by the fact that she is his love interest.

4.4 Analysis of Comic Book cost

Over the period from 2001 to 2016, we observed that Iron Man had a higher face value as compared to Captain America for every year except 2011 and 2012. Interestingly in 2011, the face value of Captain America increased approximately 1.7 times Iron Man's face value which interestingly coincides with the release of movie, 'Captain America: The First Avenger'.

5 FUTURE DIRECTIONS

We plan to enhance our 'Analysis of Comic Book cost' visualization by providing the user with two drop downs with list of top fifty Marvel characters so that the user can identify different trends and draw interesting conclusions.

REFERENCES

[1] Johnston, R. (2016). DC Beat Marvel's Marketshare In July 2016 - 41% With A Third Fewer Comics - Bleeding Cool Comic Book, Movie, TV News.

[2] Silva, S. F., & Catarci, T. (2000). Visualization of linear time-oriented data: a survey. In *Web Information Systems Engineering, 2000. Proceedings of the First International Conference on* (Vol. 1, pp. 310-319). IEEE.

[3] Lu, A., & Shen, H. W. (2008, March). Interactive storyboard for overall time-varying data visualization. In *2008 IEEE Pacific Visualization Symposium* (pp. 143-150). IEEE.

[4] Alberich, R., Miro-Julia, J., & Rosselló, F. (2002). Marvel Universe looks almost like a real social network. *arXiv preprint cond-mat/0202174*.

[5] Harris, C. (n.d.). Visualizing the Marvel Comic Universe.

[6] Tim Leong, *Super Graphic: A Visual Guide to the Comic Book Universe*, Chronicle Books LLC, 2013. Retrieved from https://books.google.com/

[7] Ardanuy, M. C., & Sporleder, C. (2015). Clustering of Novels Represented as Social Networks. *LiLT (Linguistic Issues in Language Technology)*, *12*.

[8] Bemong, N., Borghart, P., De Dobbeleer, M., Demoen, K., De Temmerman, K., & Keunen, B. (2010). *Bakhtin's theory of the literary chronotope: Reflections, applications, perspectives* (p. 213). Academia Press.

[9] Andrews, D., & Baber, C. (2014, April). Visualizing interactive narratives: employing a branching comic to tell a story and show its readings. In*Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (pp. 1895-1904). ACM.

[10] Diakopoulos, N., Kivran-Swaine, F., & Naaman, M. (2011, May). Playable data: characterizing the design space of game-y infographics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1717-1726). ACM.

[11] Agarwal, A., Corvalan, A., Jensen, J., & Rambow, O. (2012, June). Social network analysis of alice in wonderland. In *Workshop on Computational Linguistics for Literature* (pp. 88-96).

[12] Henry, N., Fekete, J. D., & McGuffin, M. J. (2007). NodeTrix: a hybrid visualization of social networks. *IEEE transactions on visualization and computer graphics*, *13*(6), 1302-1309.

[13] Heer, J., & Shneiderman, B. (2012). Interactive dynamics for visual analysis.*Queue*, *10*(2), 30.

[14] Silvia, S., Abbas, J., Huskey, S., & Weaver, C. Storyline Visualization with Force Directed Layout.

[15] Bostock, M. (n.d.). Les Misérables Co-occurrence.

[16] Talon, D. S. (2007). The Technology and Storytelling of Color in the Sequential Narrative. *International Journal of the Book*, *4*(4).

[17] Lim, S., & Seodaemun-gu, S. Data visualization design for exploring usage tendency according to character identity-focused on Marvel comics characters.

[18] Marvel Chronology Project - Main. (n.d.).