Exploring Temporal Network Patterns and Character Interactions in the Naruto Series

Mona Saleem[a], **Ata Canpolat**[b], **and Lukas Olsen**[c]

This manuscript was compiled on December 6, 2023

This research paper explores the dynamic social networks within the Japanese animated series "Naruto" using network analysis tools. The series span over 1013 episodes, covering three seasons and involving 39 villages and 1037 characters with distinct attributes. The objective is to systematically identify the changing importance of characters over time and explore possibilities of underlying important characters. This is achieved by treating the series as evolving dynamic co-occurrence networks. The networks are undirected and weighted, reflecting the frequency of interactions between characters. Data is fetched from the Naruto fandom page, providing episode text and character information. By applying network analysis tools such as centrality measures, community detection algorithms, sentiment analysis, and word clouds, we identify key figures. Results showcase the evolving complexity of character interactions, with fluctuating power law exponents indicating changing dynamics. Centrality measures highlight key characters, while community detection reveals the impact of affiliations like villages on narrative development. Word clouds align with degree centrality, offering a holistic view of character roles and relationships. While acknowledging limitations in edge weight subjectivity and community detection sensitivity, the paper suggests future research directions, including dynamic edge weight thresholds and refined attribute-based partitioning. The study concludes by emphasizing the broader implications for understanding character evolution in complex, dynamic fictional universes.

Network Science | Social Dynamics | Character Interactions | Naruto Series | Community Detection | Sentiment Analysis

Network science has since its beginning in Kaliningrad (1) proven to be a powerful tool in regards to examining dynamics and structures of social systems. Thus, we found interest in delving deep into the social dynamics in the Japanese animated series *'Naruto'* by Masashi Kishimoto (2). In the series, Naruto Uzumaki is presumably the main character for the initial two seasons accompanied by his friend and primary arch-nemesis, Sasuke Uchiha. The narrative in the series shifts towards the final season, focusing on Naruto's son *Boruto*. The abundance of characters presented in *Naruto* have varying roles and popularity throughout the series, begging the question of how the series and the characters' importance evolve over time.

The research seeks to answer this through a social network analysis by utilizing selected network tools. The *Results* section of this paper will describe in detail the findings upon conducting network analysis and show how quickly the graphs become dense with surprisingly different changes in character importance and network structure. Furthermore, we showcase the networks for interactions between the characters at different threshold levels based on the frequency of appearances and word-clouds displaying the development of the most prominent characters and words in each season of the series. Lastly, we show the partitioning of the characters by using attributes such as villages and sentiment scores in a community detection analysis.

Results

Social interactions can be constructed as different networks given the research to be conducted with each network displaying different dynamics and properties. This naturally

Significance Statement

This research employs network analysis tools to uncover the dynamic social interactions within the extensive narrative of the Japanese animated series "Naruto." Spanning over 1013 episodes, the study delves into the evolving importance of characters, systematically analyzing undirected and weighted networks that reflect the frequency of interactions. By uncovering changing power dynamics and character affiliations, the research not only enhances our understanding of complex fictional universes but also showcases the potential of network science in analyzing and exploring narrative evolutions for large animated series.

Author affiliations: [a]Mona Saleem; [b]Ata Canpolat; [c]Lukas Olsen

Author contributions: M.S. & A.C. did the network analysis. A.C. & L.O. did the data fetching. L.O. did the textual analysis. M.S., A.C., & L.O. wrote the paper.

influences the choice made of the given 1013 evolving co-occurrence networks of the Naruto characters. The chosen construction of the networks is undirected and weighted graphs with a weight of 1 added for each time two characters co-occur in an episode.

As a result of this, the networks are remarkably dense from the initial episode and continue to be as new characters are introduced (Fig.1A). Therefore the network measures and properties are also dynamic and prove to impact the popularity and centrality of key figures in the series.

Degree distribution as a measure for character interaction. Character interactions are in the network equivalent to co-occurrence and links. Hence the degree distribution is a relevant aspect to examine. In terms of scale-free networks(3), the resulting graph of the last episode follows a power law distribution with a power law exponent $\gamma = 2.14$. This is a strong indicator that there are important characters established at the end, but the journey there will be shown to be competitive among the characters. Despite the low power law exponent at the final network, scale-freeness is not a constant. Due to the dynamic nature of the evolving graphs, a number of 476 times, does it occur that the graphs deter from following a power law distribution (Fig.1B). Instead, the networks are observed to be in the random regime 266 times with $\gamma > 3$ and surprisingly in the anomalous regime 210 times with $\gamma < 2$ (Fig.1C). However, regardless of the observed anomalous regime, the networks did seemingly not exhibit odd behavior upon further inspection as otherwise anticipated. This may be due to the fact that the number of characters is not sufficiently high for any fatal network consequences (4).

The shifting power law exponent at the beginning and in the middle of the series may indicate a shift in character interactions. This is mostly observed in season 1, which may be due to the early onset of introducing many characters and possibly the many filler episodes Naruto is known to have, such that these characters only interact with a selected few others for the specific filler. There is a peak again in the middle of the series around episode 482, where the Great Shinobi War officially starts (Fig.1C). This is the longest arc in the series, where many already introduced characters re-appear along with new ones, only to reveal important roles as Obito Uchiha among many others. The changing power law exponent possibly sheds light upon flux central characters which is further examined with centrality measures.

Most important characters - centrality measures. In order to analyze the development of prominent and important characters, there have been selected 9 characters that are hypothesized to play an influential role in the series listed in figure 2. All three degree-, eigenvector- and betweenness centrality measures have been applied to all the networks in an attempt to identify the important characters with 2 showing the degree centrality. All three measures overall resulted in the same patterns for the characters. From the measures exclusively, one can unsurprisingly conclude that Naruto is at all times the most popular character. However, Naruto's centrality declines over time with an onset at around episode 300. Naruto's decline correlates with the increasing centrality of the other characters. As mentioned previously, Obito Uchiha, along with his companion Madara Uchiha, are

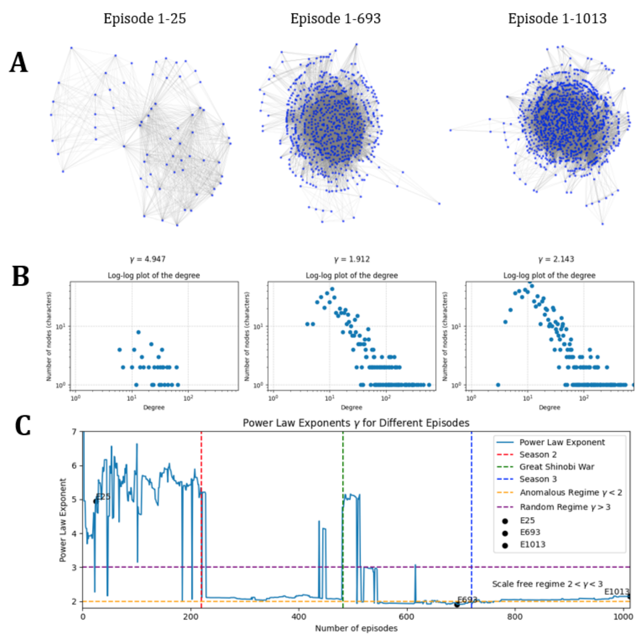

Fig. 1. Development of the co-occurrence networks of Naruto characters. (A) Visualization of the undirected evolving networks from 3 selected time points. In episodes 1-25, the network is seen to already be dense and a giant connected component. New characters continuously appear and interact with each other growing the density of the network from episodes 1-693 and the resulting network of all episodes where links are hardly distinguishable due to size and heavy interactions. (B) The log-log plots of the degree distribution from the selected time points showcase the dynamics of the character interactions. The 3 networks are seen to be in 3 different regimes at different times, respectively in the random regime with $\gamma = 4.947$, in the anomalous regime with $\gamma = 1.912$, and lastly in the scale-free regime with $\gamma = 2.143$ where the latter two have a long heavy tail. The low power law exponent indicates that popular characters indeed exist interacting with a significant number of the other characters. (C) The graph of the power law exponent γ for the networks throughout the entirety of the series. Timestamps have been visualized by respectively red and blue dotted lines for season 2, season 3, and green for the pivotal moment of the Great Shinobi War along with the regime thresholds and the networks in (A). Clear to observe that the networks undergo changes in the first and second seasons resulting in seemingly more random interactions between characters.

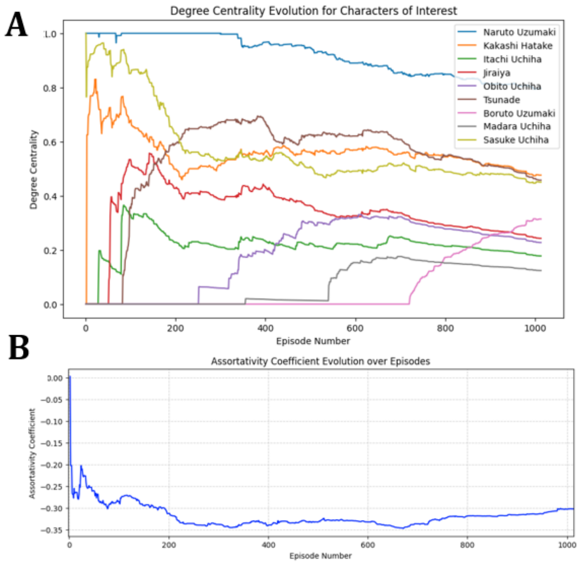

Fig. 2. (A) The dynamic degree centrality for 9 prominent selected characters. Naruto is the most central character at all times. Steep increases in centralities are observed for Tsunade and Obito Uchiha as Tsunade becomes the leader of Naruto's village and Obito proving to be an emerging strong villain. The two succeed the centrality of other characters challenging the shifting dynamics of centrality. As other characters become more popular, Naruto's centrality decreases. Evident discrimination between season 3 and the first 2 seasons by the rapid centrality increase of Boruto. (B) The degree assortativity coefficient throughout all the episodes. It remains negative at all times except for the first episode. Showcasing a slightly disassortative behavior for all networks with the coefficient ranging from [-0.2, -0.35] indicating the tendency that popular characters interact with less popular characters

seen to be having a fairly steep increase in centrality as it is revealed around this time of the series, that they are the two masterminds behind the Great Shinobi War becoming Naruto's enemies. This shift in centrality aligns with the story plot in the series that climaxes between Naruto battling Obito and Madara Uchiha who proves to be suitable strong and important opponents for the outcome of the war. These 2 characters are introduced briefly seemingly unimportant, only to become central as the series progresses.

A surprising result is the observation of the character Tsunade being the second most central character for an entirety of over 400 episodes. She is arguably not considered the main character however the degree of centrality represents that she frequently co-occurs with the other characters. The most surprising observation is of the character Sasuke Uchiha being succeeded in centrality by his fellow ninjas from around episode 200. This is a fairly staggering result taking into consideration, that Sasuke is the arch-nemesis of Naruto with the first 2 seasons maintaining an underlying theme of Naruto turning his rival-turned-friend on his side again. However, upon further inspection, the findings concluded that Naruto appears 450 times without Sasuke yet Sasuke only appears 51 times without Naruto indicating that his appearance may rely on Naruto. The onset of season 3 revolving around Naruto's son Boruto, showcases how he makes a steep and quick acceleration in centrality succeeding 4 characters as the narrative shifts over to a new generation. Regardless, Naruto maintains the most central character throughout all seasons.

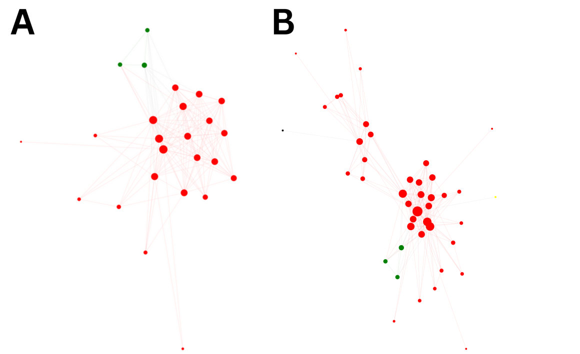

Fig. 3. (A) Network visualization of episodes 1-220; the entirety of season 1. (B) Episodes 1-1013; the whole series. Both networks are subject to a minimum co-occurrence constraint of 100. The nodes are colored w.r.t. to the village affiliations of the characters, red for Konohagakure where Naruto belongs and green for Sunagakure. While the connections between central characters get denser with time, and indicate Naruto's evolution into a central hub, the emerging clusters highlight distinctions between seasons and narrative shifts in each era.

A. Disassortativity. An intriguing aspect of the network is the pattern and tendencies of the interactions between the characters. To inspect this, the measure of the degree assortativity coefficient has been analyzed dynamically. Disregarding the first episode, the degree assortativity coefficient remains negative at all times within the range [-0.20, -0.35] showing disassortativity. This interestingly implies, that the popular characters in terms of degrees, do not solely interact with each other. Rather, it implies that many side characters of low degrees and hence appearing a few times, most likely co-occur with popular characters. This may very likely be due to the combined facts, that the network is at all times a giant connected component with a power law exponent γ being 70% of the time below 3. In other words, there are few popular high degrees characters and many unpopular low degrees characters but they all remain connected in the same network leading to disassortativity. This further suggests a diverse range of character interactions and well-mixed characters in the Naruto Universe.

Network Development with minimum Co-occurrence constraint. To provide a comprehensive overview of the evolving social dynamics of characters, a filtration is made. An edge weight threshold of 100 is applied to the character co-occurrence networks. The nodes are colored according to the villages, with the most characters coming from the villages *Konohagakure* and *Sunagakure*. This intentional filtering aims to highlight important interactions by filtering out the less frequent characters. The network analysis reveals Naruto's position as a central hub, implying his evolving significance in the series. As the narrative progresses, Naruto's interactions become more pronounced, reflecting his journey as a central figure in the series.

The series' antagonists emerge as nodes in the filtered networks and therefore illustrate their importance. Among these examples are Sasuke Uchiha and Orochimaru. The co-existence of both protagonists and antagonists for a high edge weight shows that the focal point of the series is not solely on the protagonist but rather on both figures.

Fig. 4. The upper word-cloud is made from season 3, and the lower word-cloud is made from season 2. The word-cloud made from season 1 has been left out since it has a lot of similarities with the word-cloud of the second season. The word-clouds give a good indication of the most prominent characters for each season in the Naruto series and are a nice way of showcasing the most occurring words from the text analysis of each episode.

A distinctive shift is observed when transitioning to the "Boruto" series, with Naruto's son taking center stage. The networks reflect this change, with an observation of a distinct clustering starting to emerge which primarily contains the newly introduced characters in season 3 (Fig.3B) indicating a shift in character importance.

Word-clouds. Included in the NLP part (5) of this paper, word-clouds were constructed for all three Naruto seasons to give an indication of the most prominent characters as the series developed. This was done by fetching a text description for each of the 1013 episodes. Then the text was processed and cleaned using the *NLTK* (6) library in Python. After removal of stop-words and other unwanted text, the words were tokenized and the word-clouds were computed using the Python package *WordCloud*.

When comparing the words in the word-clouds (Fig.4) to the degree centrality (Fig.2A), interesting findings come across. The word-clouds and the degree centrality seemingly align with each other. In the last season, the word-cloud showcases that the character *Boruto* is the character that occurs the most in the episode descriptions, whilst the degree centrality also displays his prominent role in the series when he is introduced in the last season. Another interesting find is in the word-cloud for season two. The character *Tsunade* is mentioned less than the characters *Jiraiya*, *Kakashi*, *Sasuke* & *Itachi*, although she is the second most central character according to degree centrality (Fig. 2.A). This goes to show that she might appear in many episodes and thereby link to many of the characters, but that her role is not as influential as the others to be worth mentioning.

Sentiment Analysis. Sentiment scores of the characters were computed based on the text that was fetched from the character descriptions. For this purpose, the *LabMT* (7) package provided a word list from which the sentiment scores was computed. After prepossessing the raw character text, the sentiment score of each character was computed. The range of the sentiment scores went from 2.49 to 4.68, making for an absolute difference of 2.19. It might be a little surprising that none of the characters is above the neutral score of 5, but given the nature of the series, it makes sense since a lot of the character descriptions will contain words such as fighting, killing, and other words that relate to a fighting ninja universe. Finally, each character sentiment score was added as a node attribute.

Community Detection. In the series, characters are parts of different villages. The different villages characters belong to can contribute to influencing factors in terms of the importance of the characters as seen in the threshold network with characters mainly belonging to Konohagakure or Sunagakure. By applying graph partitioning methods throughout the episodes, it was to be answered whether the village affiliation of the characters is the most crucial factor in how the story unfolds with time, or whether any other latent factors influence the development of the series. A community analysis for character co-occurrences was performed, employing a threshold of 100 to focus on the most significant connections throughout the series. In the following, the results were presented, focusing on the network structure in episodes 20, 100, 720, and 1013, and exploring the modularity scores to compare community structures found by manual graph partitioning based on the affiliation of the characters, and algorithmic community detection by using the Louvain algorithm (8).

Analysis of Louvain Algorithm vs. Manual Partitioning. As seen in Table 1, the modularity scores obtained through the Louvain algorithm consistently outperformed those achieved through manual partitioning based on character affiliations. This discrepancy suggests that the Louvain algorithm captures community structures that align more effectively with dynamic character interactions rather than static affiliations to specific villages. The higher modularity implies a more cohesive and internally connected community structure within the networks detected by the algorithm.

Temporal Evolution of Communities. An intriguing observation emerged from the exploration of the temporal evolution of communities. The modularity score of the network for episode 1013, the final episode, exceeded the modularity score of episode 720, the final episode of season 2, by more than twofold. This suggests a progressive tightening of community structures and character relationships toward the resolution of the series. This observation is also supported by the visualization of the partitioning as seen in Figure 5. Furthermore, when looking at the characters in the communities for each of the episodes that were analyzed, it is observed that the most important characters that are densely connected tend to stick in the same partitioning as the series evolves. However, the characters of less significant importance tend to be clustered into different partitions at different parts of the storyline.

Discussion

Limitations. While we believe that our analysis provides a comprehensive view of the Naruto series, it is crucial to recognize certain limitations that may impact the interpretation of our findings.

Table 1. Comparison of the modularity scores of Louvain algorithm and attribute-based manual partitioning for different episodes

Cumulative Networks	Village-based Partitioning	Louvain Algorithm
Episode 20	0.1226	0.1276
Episode 100	0.0451	0.1451
Episode 220	0.0418	0.1247
Episode 720	0.0262	0.1366
Episode 1013	0.2293	0.2827

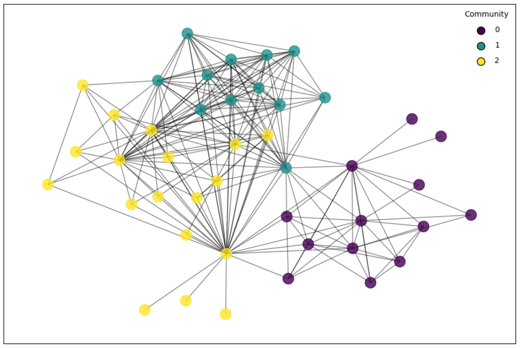

Fig. 5. Network visualization of the last episode of the series with a minimum co-occurrence threshold of 100. Coloring of the nodes shows the communities found by the Louvain algorithm. The modularity score of the partitioning is 0.2827. Three communities are detected by the Louvain algorithm for the cumulative network of the last episode, where one of the communities is easily distinguishable from the other two. Further analysis of these communities showcases that the nodes found in Community 2 are the nodes that have been introduced or predominantly appear in season 3, whereas the nodes in Communities 0 & 1 consist of the important characters from Season 1 & 2.

Considerations About Network Construction: The most important consideration revolves around the choice of the construction of the networks. Basing the networks on episode texts and not transcriptions shapes the nature of the network to have possible loss of information. This is due to the assumption that the characters listed in an episode interact with each other which may in reality not be the case. Characters could very likely appear in the same episodes in separate scenes, hence the links and all the networks may not be as accurate as desired.

Attribute-Based Partitioning Challenges: The manual partitioning of the network based on character attributes, such as village affiliation, introduces its own set of challenges. Ambiguities in character allegiance or the lack of observation of latent variables may affect the accuracy of the partitioning.

Sentiment Analysis Complexity: The sentiment analysis captures the emotional tone of both the characters and the episodes. However, the neutral use of language in synopsis of characters and episodes limits the precision of sentiment scores and their interpretation, both on the character level and also on the episode level.

Future Work. As we conclude our analysis, we mention opportunities for future research that may possibly help to uncover deeper layers of the narrative beyond the scope of our research.

Episode transcriptions: For future considerations, one could consider combining the network with textual analysis of the transcriptions of each episode. This may deliver more interesting insights about the character interactions and evolution of the series that are otherwise not detectable with solely episodes synopses.

Dynamic Edge Weight Thresholding: Future research could explore dynamic edge weight thresholds that adapt to the evolving narrative of the series. This approach may provide a more nuanced understanding of character significance over different story arcs.

Attribute-Based Partitioning Refinement: Improving the accuracy of attribute-based partitioning involves refining character attribute annotations and considering additional attributes such as character roles or power levels. This could lead to more accurate community assignments.

Fine-Tuning Sentiment Analysis: Developing a more sophisticated sentiment analysis model that considers context and character relationships may yield richer insights into the emotional dynamics of the series. Neural networks for sequential data or other statistical learning algorithms for NLP can be employed to estimate more accurate sentiment results.

Materials and Methods

The Dataset. As the main data source, the Naruto Fandom Wiki webpage is used (9). All the data about the characters, episodes, and related information has been fetched from the respective pages on Naruto Fandom Wiki.

Data Fetching. The data for this research paper was gathered using the Python package *BeautifulSoup*. BeautifulSoup is a well-known package for pulling data out of an HTML structure. Therefore, this was used when fetching data from the Naruto fandom page. Furthermore, the *Asyncio* library was also utilized in the fetching part in order to fetch the thousands of data point that was required.

1. AL Barabási, Network science (n.d.) Chapter 2, Accessed: 6 December 2023.
2. Narutopedia (n.d.) Accessed: 6 December 2023.
3. AL Barabási, Network science (n.d.) Chapter 4, Accessed: 6 December 2023.
4. AL Barabási, Network science (n.d.) Chapter 7, Accessed: 6 December 2023.
5. S Bird, E Klein, E Loper, *Natural language processing with Python: analyzing text with the natural language toolkit.* (" O'Reilly Media, Inc."), (2009).
6. Natural language toolkit (https://www.nltk.org) (2023).
7. Labmt (https://labmt-simple.readthedocs.io/getting-started.html) (2023).
8. VD Blondel, JL Guillaume, R Lambiotte, E Lefebvre, Fast unfolding of communities in large networks. *J. statistical mechanics: theory experiment* **2008**, P10008 (2008).
9. Narutopedia (https://naruto.fandom.com/wiki/Narutopedia) (year?).