

Network Analysis of the Matrix Trilogy

Authors: Nataliia Romanenko, Cole Johnson, and Sawyer Marchand

Network science is an interdisciplinary and growing field that studies complex networks by combining graph theory and data science to gain insight into the connections and relationships around us. Network analysis of character interactions in a movie can be used to answer questions about the social structure of the movie, and to identify different communities of characters, social network characteristics, and relationships. The purpose of our project was to use network science to analyze the social structure and social relations in the Matrix trilogy: *The Matrix* (1999), *The Matrix Reloaded* (2003), and *The Matrix Revolutions* (2003). Our research project included all aspects of the network analysis process: data collection, network construction, visualization, and analysis. We used the programming language R for data collection, data cleaning, statistical and sentiment analysis; and the network analysis software package *Gephi* for data modeling and visualization. Based on character interactions that were extracted from movie scripts, we built a character interaction network for each movie and analyzed them by calculating different network measures such as weighted degree centrality and betweenness centrality. In addition, we identified the top six characters in each movie in the Matrix trilogy and performed a sentiment analysis. We concluded that the most important and the most connected character, based on weighted degree centrality and betweenness centrality, in all three movies was Neo. Furthermore, we used modularity-based community detection to determine the community structure for the movies. This analysis enabled us to better understand the social network structure and social dynamics across all three movies.