12-2014

Analyzing Shakespeare's Plays in a Network Perspective

Vikas Thotakuri
University of Nebraska at Omaha

Analyzing Shakespeare’s Plays in a Network Perspective

A Thesis

Presented to the

Department of Computer Science

And the

Faculty of the Graduate College

University of Nebraska

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science

University of Nebraska at Omaha

By

Vikas Thotakuri

December, 2014

Supervisory Committee:

Dr. Sanjukta Bhowmick

Dr. Yuliya Lierler

Dr. Robin A. Gandhi

Analyzing Shakespeare's Plays in a Network Perspective

Vikas Thotakuri

University of Nebraska, 2014

Advisor: Dr.Sanjukta Bhowmick

Abstract

Networks are popular models for representing interactions between entities in systems, such as in sociology, bioinformatics, and epidemiology. The entities in the networks are represented as vertices and their pair-wise interactions are represented as edges [1]. Many network metrics such as degree centrality (number of connections of an entity) and betweenness centrality (number of shortest paths passing through the entity) have been developed to rank the entities according to their importance [7] [10].Social networks are generally modeled on only one type of relation. Groups are open-ended, which means the number of participants and the time frame are not finite. Time frame may not cover significant events and their effect. How would the analysis change if we modeled the interactions and relationships of a closed group, over significant incidents? It is difficult to obtain real life data, because of the time commitment and privacy constraints. The next best option: Analyze fiction, which would give an indication of social relations [15].In this thesis, we study the effectiveness of these metrics in closed-form social interactions—particularly in the context of Shakespeare's dramas [2] [18]. In plays the dialogues amongst characters are very precise to express the gist of their interactions in a short time frame. We are interested in understanding how this sort of interaction differs in

a qualitative sense from the interactions seen in social media such as Facebook and Twitter. Our observations show that the popular network metrics are not always successful in correctly identifying the lead characters of the play and we propose a new method of creating two different types of networks from each play by using different criterion. Also the third type of network model, the Time Series Analysis considers the important characters of each play and filters edge lists based on them. Here the occurrence/influence of the important characters is examined from scene to scene and in turn from act to act from beginning to the end of each play we considered.

Acknowledgement

Foremost, I am extremely thankful and indebted to my advisor Dr. Sanjukta Bhowmick, who encouraged me to take the rewarding path in completion of my thesis. Besides my advisor, I would like to thank the rest of my committee: Dr. Yuliya Lierler and Dr. Robin A. Gandhi for their guidance. My sincere thanks to my graduate advisor Carla Frakes, who help me complete the paperwork and continuously encourage me to meet the deadlines. My heartfelt thanks goes to my parents Srinivasa Rao, Vijaya laxmi and my sister Vineela who have motivated me continuously, in every possible way, from halfway around the world.

I would also like to thank my grandparents, cousins and friends for their extensive support in all possible ways.

Last but not least, I would like to thank the college of IS & T for providing me the great infrastructure to work on my thesis.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Data is complex to analyze. In order to advance and observe the data in a keen manner, it should be represented in an organized format. One of the popular ways to represent data is in the form of networks, where entities are considered as vertices and the relationships constitute edges (for example, Social Networks, Biological networks) [1]. However raw data needs to be processed, cleaned and modelled in order to represent the desired data as a network. In this Thesis, we emphasize on Shakespeare's Plays which are best examples of complex unstructured data, because the text in the play is not organized and doesn't follow a certain pattern. They are a large collection of data where understanding them becomes a complex task. As the data is unfiltered, initially we need to figure what aspects in the data should be considered in order to model the network and analyze it. Hence, we do the data filtering based on the list of characters/people and their roles in the play.

Plays, stories can be represented as networks [3]. The list of people in a play can be modeled as vertices and communication/role as edges in the network [1]. The reason we considered plays is, they are a closed group of characters with finite amount of conversation, and the plays are available in text online.

Apart from this, if we consider a social network for example, Facebook or Twitter it is more of a personal conversation/messages between people where the privacy constraint comes into picture. Social network's data change from time to time which is difficult to track and verify.

Several researchers have applied graph theory on social networks, where they are concentrated on a particular event or time frame. For example, Twitter analysis is based on the hash tags and is done based on an important event or person (for example fifa2014 or Obama) [15]. As a result they cannot track the whole story as it doesn't have accurate information as everybody tweets in their own way and different languages. It becomes very complex.

In this thesis, we considered a set of Shakespeare's plays, represent them in the form of networks and analyze them.

We study the effectiveness of different network metrics like closeness centrality, betweenness centrality, eigenvector centrality (which we discuss later) in closed-form social interactions. In Shakespeare's plays the dialogues between characters are very precise to express the gist of their interactions in a short time frame [18]. For example, a complete play will last for three to four hours when performed. We are interested in understanding how this sort of interaction differs in a qualitative sense from the interactions seen in social media such as Facebook and Twitter.

A network is called a directed network when the edges between the nodes have arcs which denote the direction of flow. Undirected networks are which doesn't have any arcs to edges and an edge is considered as bi-directional always.

We design three different types of networks 1) **Interaction** networks, connect two characters when they appear in the same scene and the edge weight is the number of lines spoken. These are undirected networks. 2) **Mentioning** networks connect two characters if one is mentioned by the other. Edge weight is the number of mentions. These are directed networks. 3) **Mentioning with relationships**, connect two characters, character with a relationship (For example Father, Mother, Brother) when the character mentions the other character with name or the relationship. Edge weight is the number of mentions and the network is directed.

After the networks are designed we visualize them using a network visualizing tool Cytoscape [5]. Further metrics are computed for both interaction and mentioning networks considering the list of important characters/people in each play using Gephi [6].

Then we examine the occurrence of important characters in each scene and their role in the play as we go scene by scene; by this we can find the role and influence of the important characters in the play, which we termed as Time Series Analysis.

In this thesis we concentrated on women centric/heroine oriented plays. The reason is, when we consider woman centric plays, the results show up differently than expected which we show as the difference between the two models interaction and mentioning.

For the important characters in a play, as Shakespeare's plays are very well known we know the important characters like hero, heroine. There are multiple websites (One of the website: http://www.sparknotes.com/shakespeare/)which list the important characters of each play.

1.1 Contribution

- We have collected open-source formatted text from MIT’s website [18].Filtered the plays based on the ACT’s and SCENE’s.
- Research on different ways of creating multiple networks using single dataset.
- We have designed and implemented algorithms for extracting both interaction networks and mentioning networks from each play considered.

- Visualization and computation of different metrics for the networks generated.
- Worked on creating networks based on different criteria which are named as Character by Character and Full Scene.

In Character by Character criteria, when a character talks we consider that he/she is talking to the immediate next character that is going to appear in the scene.

In Full Scene criteria, when a character talks we consider he/she is talking to all the characters present in the whole scene.

- We have created a web tool which is used to read the plays and create edge lists/networks for both interaction and mentioning.

1.2 Outline of Thesis

The thesis is organized as follows. In Chapter 2, we discuss about the background of networks and graphs and brief our application. In Chapter 3, we discuss the implementation details like creation of model, data extraction, how to create relationships, issues and static analysis [15]. In Chapter 4, we discuss implementation of Time series analysis/links. In Chapter 5, we talk about the web tool and its working. In Chapter 6, we conclude our thesis and discuss about the present potential ideas about the future research. Chapter 7 is the appendix where we have all the results from the Gephi analysis and Cytoscape.

Chapter 2

Background

A real-world dataset can be easily analyzed by representing it in the form of a graph/network (for example Social Network). A graph can be defined as the collection of objects which are identified as vertices/nodes connected with links which are termed as edges in the graph theory. A graph is a set of vertices connected by edges [1]. Graphs are extensively used in the field of mathematics and computer science. For example in social network analysis, people are considered as vertices and communication between them is represented as an edge. We consider multiple network properties/metrics in order to analyze the graphs. The properties are classified into two categories a) vertex based properties and b) network based properties.

2.1 Graph Terminology

A graph is collection of vertices and edges. Formally, G = (V, E) consists of set of vertices V and edges E, where E is subset of (V x V). There are two types of graphs i) directed and ii) undirected. A graph is directed if edges point in one direction from one vertex to another vertex, otherwise a graph is undirected.

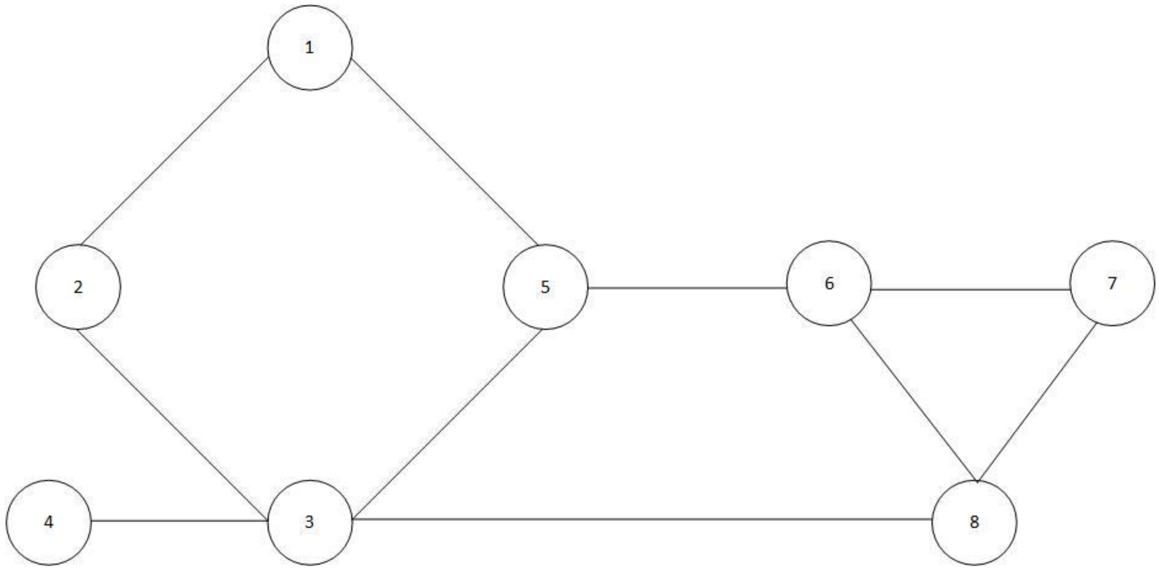


Figure 1: 2.1 Undirected Graph

2.2 Graph Properties

2.2.1 **Vertex Based Properties**

Vertex based properties are defined per vertex of the network. Some of them are

2.2.1.1 Degree

The degree of a vertex in a graph is the number of edges the vertex shares with the other vertices [7]. The degree of vertex v is denoted as deg(v). In figure 2.1, degree of vertices are deg(V1) = 2; deg(V2) = 2 ; deg(V3) = 4 ; deg(V4) = 1 ; deg(V5) = 3 ; deg(V6) = 3 ; deg(V7) = 2 ; deg(V8) = 3.

2.2.1.2 Betweenness Centrality

Most of the shortest paths in a network go through the vertices with the high betweenness centrality [8]. Therefore, these become more the central point controlling the communication. Betweenness Centrality of a vertex v is calculated as sum of the ratio of the number of shortest paths in the graph include vertex v to the total number of shortest paths in the graph. The betweenness centrality BC (v) of a vertex v ∈ V is the sum over all pairs of vertices u, w ∈ V, of the fraction of shortest paths between u and w that pass through v

$$BC(v) = \sum_{\substack{u,w \in V \\ u \neq w \neq v}} \frac{\sigma_{u\,w}(v)}{\sigma_{u\,w}}$$

Where$\sigma_{u\,w}(v)$ denotes the total number of shortest paths between u and w that pass through vertex v and $\sigma_{u\,w}$ denotes the total number of shortest paths between u and w.

2.2.1.3 Eigenvector Centrality

Eigenvector centrality measures the importance and influence of a node on others in the network. Vertices with high Eigenvector scores have many connections and their connections have many connections [7].

2.2.1.4 In degree

In a directed network, the number of edges coming into a vertex is defined as its in-degree [9].

2.2.1.5 Out degree

In a directed network, the number of edges travelling away from a vertex is defined as its out-degree [9].

2.2.1.6 Page Rank

Page rank is a measure used to determine the importance of a node/vertex. It is computed by a rough estimate of how many edges traverse from or to the vertex. Page rank is calculated based on the in-degree and out-degree of a vertex. Google uses Page Rank algorithm to rank the websites [9].

2.2.2 Network Based Properties

Network based properties are defined over entire network. Some of them are following

2.2.2.1 Vertices

The total number of vertices in a graph. There are a total of eight vertices in the graph from Figure 2.1.

2.2.2.2 Edges

The total number of edges in a graph. There are a total of 10 edges in the graph from Figure 2.1.

2.2.2.3 Degree Distribution

Degree distribution is the distribution of the different degrees (and their frequency) of the vertices over the network. Most scale free networks like social networks observe a power law distribution that is there exist many vertices with low degree and the number of vertices exponentially goes down as the degree increases [9].

2.3 Brief Outline of Our Project

The larger the data, the more complex it is to analyze it. Most real-world data can be represented as a network. Several researchers applied graph theory in studying social networks. In our application we tend to produce two different edge lists from each play applying different criterion for single data set. The reason we create different data network models is it facilitates in examining the datasets in multiple perspectives than in a single way.

When we consider plays it's a fun social network analysis as it is a real story. For example, most interesting plays/stories like Game of Thrones, Star Trek and

Shakespeare’s plays are different types of plays/stories in which people are more interested.

2.3.1. About William Shakespeare

William Shakespeare was born in Stratford-upon-Avon, UK, on April 23, 1564.William Shakespeare is a mysterious figure with regards to personal history. William Shakespeare's plays have great reputation in the English language and in Western literature. Traditionally, the 37 plays are divided into the genres of tragedy, history, comedy and tragic comedy; all the plays are translated into every major language and are performed around the world [19].

Shakespeare’s plays are one of the most extensively followed dramas around the globe, and hence, we can validate our results/outcome easily. There are many plays that can be researched which can confirm some fixed pattern of our results. Shakespeare’s plays are classified into acts which in turn are classified into scenes [17]. This helps us not only perform static analysis but also the dynamic which change from scene to scene. With the help of acts and scenes we can track the important events, characters and influences in the play.

Chapter 3

Implementation and Data Collection

3.1 Introduction

Shakespeare's plays are in an organized format in the MIT's website which is the web's first edition of complete works of William Shakespeare [18]. Each play is divided into ACTs and in turn each act is collection of multiple number of SCENEs. Each act and scene has multiple characters/people entering and exiting based on their role. Each character has a dialogue which is represented in text. Each play follows the same format where each scene has a description of the location where the scene is taking place (for example, the room in the palace, the forest ect.), and each person's/character's name is mentioned followed by their respective dialogues.

Understanding the data is an important task. However, due to large amount of data it is difficult to summarize it. Hence, we use the concept of networks. In order to understand the network evolution we account multiple network metrics.

We are interested in who is talking with whom and how long is the conversation between them, and how the important people are influencing the play. In the methodology section, we will concentrate on different network metrics.

Choosing datasets had been a complex task for us to perform this research/analysis. We have researched with different sources; went through multiple websites, scripts and books; and finally chose plays as appropriate resource as the compatible data sets.

Looking at the plays in a network perspective is probably new. It's difficult to find sources to refer and educate ourselves to go forward in the research.

Coming to Plays the main drawback is they are written in an older English language. Our main motto is to analyze the plays easily without reading through them completely. As MIT's website is the only source where we can find the complete works of Shakespeare we have to use those for the research.

In Shakespeare's plays, there are multiple issues which are tough to analyze. For example, there are keywords like ENTER and EXIT in the plays which are used to indicate the entry and exit of characters to/from the scenes and acts [17]. These are difficult to track.

Continuity in plays is one of the major difficulties faced in the analysis of plays. There will be characters coming in and going out from the plays. It's very difficult to track who are talking to whom. Hence, we assumed two different scenarios here.

1) Character by Character: Considering a character is speaking to the one who appears immediately after him/her in the play.
2) Full Scene: A character is talking to all the other characters present in the scene.

3.2 Methodology

We concentrated mainly on women centric tragedies and comedy plays. The list of plays considered are *As You Like It, Hamlet, Julius Caesar, King Lear, Macbeth, The Merchant of Venice, Much Ado About Nothing, Othello, Romeo and Juliet, Taming of the Shrew, The Tempest, Twelfth Night.*

The reason we considered mostly women centric plays, which are Comedies and Tragedies is as the plays are written in ancient times where the importance of a character cannot be assessed by the frequency of talk. For example, even though the queen is very important, she doesn’t have considerable dialogue or much talking as her message will be passed to people most of the times by a clown or court men.

There are hero/heroine/villain/hidden people and we can categorize people easily. We used two different criterions to extract three different types of networks from each play considered. They are 1) Interaction and 2) Mentioning. Interaction is the communication between people in the scene, and mentioning is tracked based on the occurrence of a character’s name in others dialogues. Based on this criterion we extracted two different types of networks namely 1) Interaction networks – in which we connect two characters if they appear in the same scene. Edge weight is the number of lines spoken. This is an undirected network. 2) Mentioning networks – is when we connect two characters if one is mentioned by the other. Edge weight is the number of mentions. This is a directed network.

The important characters in the play are assessed from the following sources

1) Wikipedia – Where the important/top characters in the play are listed.
2) Traditionally there are certain people who perform important roles, when the play is performed by which important characters are known.

3.2.1 Issues

There are many challenges faced in order to collect the data and perform data mining. Here are some

Challenge 1: *The issue is with the Mentioning where characters/people are not always mentioned with their names but will be mentioned by relationship or role or a pronoun (for example mother, father and clown). In fact they are mentioned indirectly most of the times than directly by their name.*

Solution 1: As the pronouns are difficult to track and map with the actual characters, we came up with a third type of network where we account for a particular list of relationships for each play (the list of relationships considered are different for each play) and extract a new network which is the third from a single data set. Mentioning with Relationships, the third – in which we consider the relationships (for example mother, father) with which the characters/people are mentioned and we connect character with relationship when they mention. This way we are able to account most of the mentions into our network excluding the pronouns. Edge weight is the number of mentions. This is a directed network.

Challenge 2: *Names of the characters have multiple spellings. For example Katharina in* Taming of the Shrew *is spelled in two different ways "Katharina" and "Katarina."*

Solution 2: It is difficult to track the name. As of now it is a very rare case we went in manually and changed the spelling to a single word.

Challenge 3: *Characters are mentioned with surnames. For example, Lady Macbeth is mentioned as Macbeth and hence it is a confusion of whether it is addressing Macbeth or Lady Macbeth.*

Solution 3: It is really a critical issue, which is solved only when we walk through the whole play. For time being we left this issue and assumed that Macbeth is the one who is mentioned when we find the word "Macbeth" in somebody's dialogue.

Challenge 4: *Words like page and prologue appears in the play as a character name*.

Solution 4: The way we read the play is hierarchical. We assume Each ACT contains Scene's and each scene has characters talking. So, here we coded such that we maintain a list of exceptional words where these types of words (page, prologue) are not considered as a character name and are ignored.

Challenge 5: *Sometimes there is grammar mentioned along with the name which is difficult to track. For example in* As You Like It *lord is mentioned as 'A Lord.'*

Solution 5: In this case we combine the grammar with the word and consider it as a single word. A Lord is combined and considered as "ALord."

3.2.2 Network Analysis and Visualization:

Network analysis is done on the combined edge lists (without considering the ACT and SCENE division).

3.2.2.1 Cytoscape:

Cytoscape is an online open source tool which is used to visualize the networks and integrating these networks with annotations, gene expression profiles and other state data [5]. Although Cytoscape was originally designed for biological research, it is now a general platform for complex network analysis and visualization [5].

Cytoscape is used to perform network analysis and further the network is visualized by mapping the node color to the network metric Betweenness Centrality [9] – Connecting nearly non-interacting groups of characters. The node size is mapped to Degree Centrality [9] –The number of different characters that share the scene.

The more red the node is indicates higher Betweenness Centrality. The larger the node, the higher the degree is.

The below are some of the sample pictures which display the Cytoscape visualization.

Figure 2: *3.1 As you like it - Interaction*

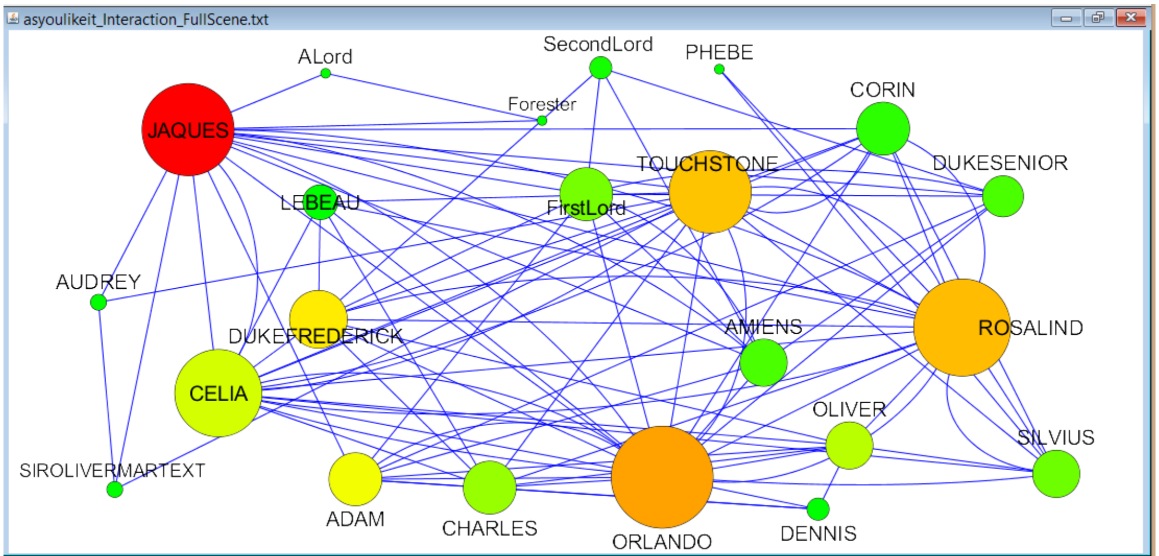


Figure 3: *3.2 Hamlet - Interaction*

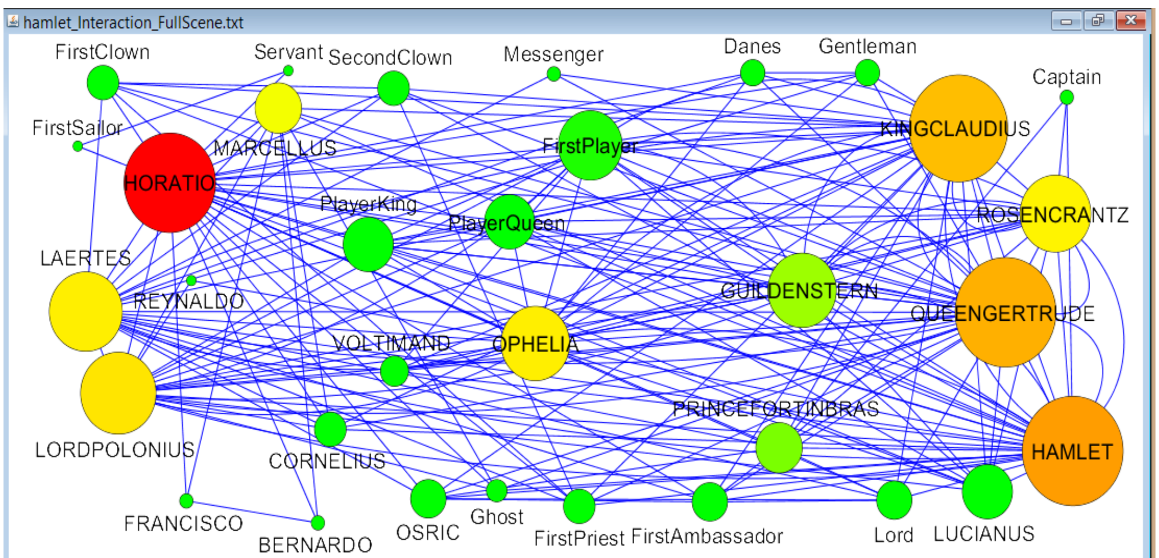


Figure 4: *3.3 As you like it - Mentioning*

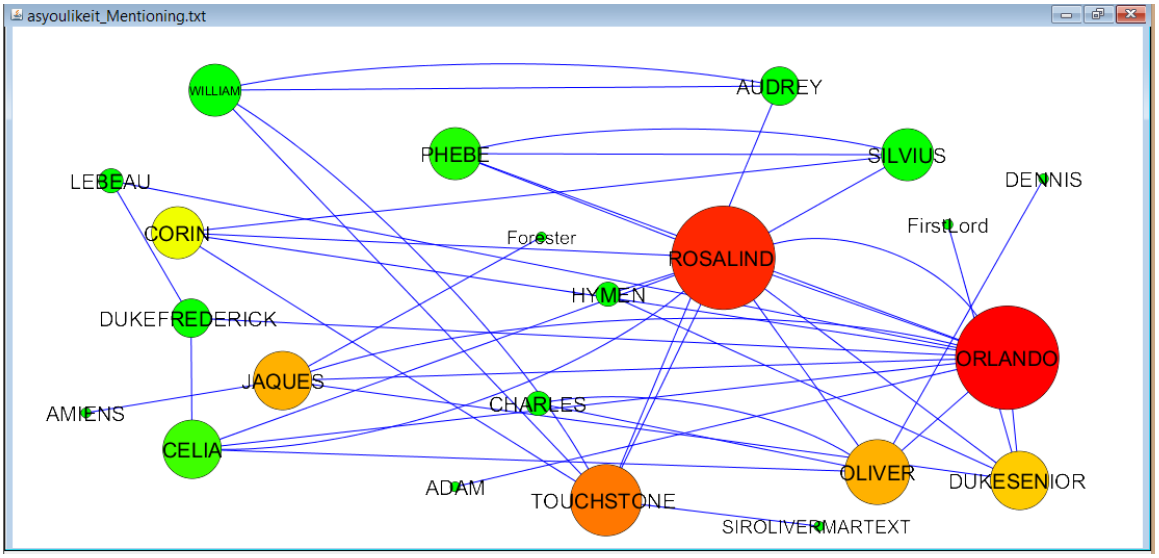
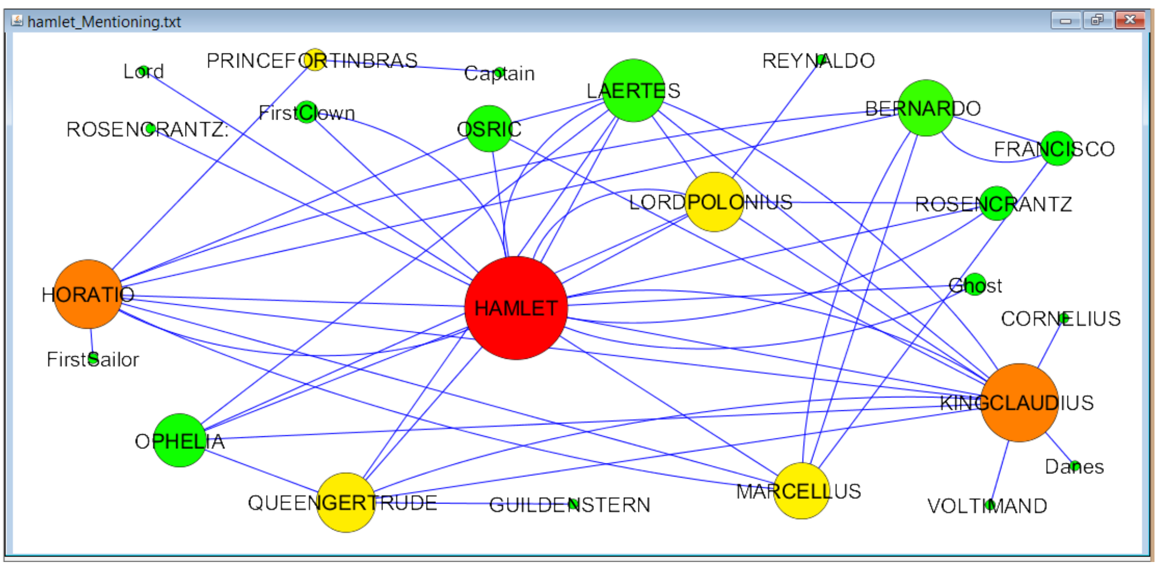


Figure 5: *3.4 Hamlet - Mentioning*



3.2.2.2 Gephi:

Gephi is an online open source tool [6]. Gephi is an exploration platform for all kinds of networks and complex systems, dynamic and hierarchical graphs [6].

Using Gephi – Network analyzing tool [6] we computed the metrics Degree, Closeness Centrality, Betweenness Centrality and Eigenvector Centrality for interaction network considering the important list of characters in each play. In-degree, out-degree and PageRank for mentioning network considering the list of important characters/people in each play. After we get the metric values using Gephi, the characters list is ranked based on the values of each metric and in turn collective average rank is calculated [14].

Initially the edge lists are imported into a Microsoft Excel sheet where the edges list is categorized into SOURCE, TARGET and WEIGHT as shown in the table below.

Table 1:*3.1 Edge List*

	A	B	C
1	SOURCE	TARGET	WEIGHT
2	ORLANDO	ADAM	199
3	ORLANDO	OLIVER	121
4	ORLANDO	DENNIS	49
5	ORLANDO	CHARLES	126
6	ADAM	OLIVER	49
7	ADAM	DENNIS	9
8	ADAM	CHARLES	36
9	OLIVER	DENNIS	42
10	OLIVER	CHARLES	85
11	DENNIS	CHARLES	33
12	CELIA	ROSALINE	629
13	CELIA	TOUCHST(	103
14	CELIA	LEBEAU	55
15	CELIA	DUKEFREE	97
16	CELIA	ORLANDO	797

Further edge lists are imported into Gephi, and network analysis is performed by running different metrics as mentioned earlier.

Here the list of people who came up based on the ranking are different from the people who are expected as we are aware of the play and important/unimportant characters in them. Now we consider the list of people who are heroes/heroines and villain in the plays and find their rankings and create a time-series analysis.

Different Formulas used to perform Gephi analysis are:

3.2.2.2.1 Rank:
RANK (number, ref, [order])

*Number**:*** The number whose rank you want to find.

Ref: An array of or a reference to a list of numbers. Nonnumeric values in ref are ignored.

Order: A number specifying how to rank *Number*. It has two values either 0 or 1.

If 0, the rank is ordered in descending order.

If 1, the rank is ordered in ascending order.

3.2.2.2.2 Average:
AVERAGE (number1, [number2]...)

Number1: Required. The first number, cell reference, or range for which you want the average.

Number2: Optional. Additional numbers, cell references or ranges for which you want the average, up to a maximum of 255.

3.2.2.2.3 Sort:
A custom sort is performed based on the Average Rank where the whole table is sorted based on the descending order of the Average Rank.

A sample Gephi result is as follows:

Table 2: *3.2 As you like it – Interaction*

1	Id	Degree	RANK	Closeness	RANK	Betweenr	RANK	Eigenvect	RANK	AVERAGE RANK
2	ORLANDO	19	22	1.333333	1	28.86825	21	1	22	16.5
3	ROSALIND	18	21	1.47619	3	23.09762	20	0.828247	21	16.25
4	CELIA	16	19	1.52381	5	7.214286	16	0.812361	20	15
5	TOUCHSTONE	15	18	1.47619	3	21.48095	19	0.792127	19	14.75
6	DUKEFREDERICK	10	17	1.666667	6	12.83571	18	0.706493	16	14.25
7	OLIVER	8	10	1.809524	11	6.325397	15	0.628659	15	12.75
8	ADAM	9	13	1.666667	6	8.321429	17	0.558579	13	12.25
9	CHARLES	9	13	1.666667	6	5.211111	14	0.72407	17	12.5
10	FirstLord	9	13	1.761905	9	4.022222	13	0.481718	11	11.5
11	DUKESENIOR	7	9	1.809524	11	2.505556	10	0.408214	8	9.5
12	AMIENS	8	10	1.809524	11	2.505556	10	0.408214	8	9.75
13	SecondLord	4	6	2.190476	19	0.666667	8	0.237265	6	9.75
14	CORIN	9	13	1.761905	9	1.492857	9	0.539961	12	10.75
15	DENNIS	4	6	2.095238	18	0	1	0.338953	7	8
16	SILVIUS	8	10	2	15	3.733333	12	0.435876	10	11.75
17	JAQUES	17	20	1.380952	2	63.71905	22	0.761236	18	15.5
18	LEBEAU	6	8	1.809524	11	0	1	0.564034	14	8.5
19	AUDREY	3	4	2.047619	16	0	1	0.208185	4	6.25
20	SIROLIVERMARTEXT	3	4	2.047619	16	0	1	0.208185	4	6.25
21	PHEBE	2	1	2.380952	22	0	1	0.147072	3	6.75
22	ALord	2	1	2.285714	20	0	1	0.104239	1	5.75
23	Forester	2	1	2.285714	20	0	1	0.104239	1	5.75

Table 3: *3.3 Take away's from Gephi Analysis*

As you like it	**Interaction**				**Mentioning**		
Characters	*Degree*	*Betweenness*	*Closeness*	*Eigenvector*	*In degree*	*Out degree*	*Page Rank*
Orlando	*High*	*High*	*Low*	*High*	*High*	*High*	*High*
Rosalind	*High*	*High*	*Low*	*High*	*High*	*High*	*High*
Celia	*High*	*Low*	*Low*	*High*	*High*	*Low*	*High*
Touch Stone	*High*	*High*	*Low*	*High*	*High*	*High*	*High*
Oliver	*High*	*Low*	*Low*	*High*	*Low*	*High*	*High*
Jaques	*High*	*High*	*High*	*High*	*High*	*Low*	*High*
Hamlet	**Interaction**				**Mentioning**		
Characters	*Degree*	*Betweenness*	*Closeness*	*Eigenvector*	*In degree*	*Out degree*	*Page Rank*
Queen Gertrude	*High*	*Low*	*Low*	*High*	*Low*	*High*	*Low*

Hamlet	*High*	*Low*	*Low*	*High*	*High*	*High*	*High*
Horatio	*High*	*High*	*Low*	*High*	*Low*	*High*	*High*
King Claudius	*High*	*Low*	*Low*	*High*	*High*	*High*	*High*
Laertes	*High*	*Low*	*Low*	*High*	*High*	*Low*	*High*
Bernardo	*Low*	*Low*	*High*	*Low*	*Low*	*Low*	*Low*

A character is marked as *High* if the value of the particular metric is greater than or equal to the half of the highest value of that metric in the table and *Low* otherwise.

From the above analysis, we can observe that Jaques from *As you like it* appear high in Degree, Betweenness Centrality, Closeness Centrality and Eigenvector Centrality and Horatio from *Hamlet* appears high in Degree, Betweenness and Eigenvector Centrality in interaction.

In case of mentioning, Orlando, Rosalind from *As you like it* and Hamlet appear high in in degree, out degree and page rank.

From this analysis, interaction shows social structure and mentioning shows the story structure.

The Gephi summary for the rest of the plays can be found in the appendix section.

Table 4:*3.4 Summery of Important Characters- Gephi Analysis*

List Of Plays	Interaction	Mentioning	Summary
As You Like It	*Rosalind, Celia*	*Rosalind*	**Romantic, Imbalanced**
Hamlet	*Queen Gertrude*	--	**Action, Balanced**
Julius Caesar	--	--	**Action, Imbalanced**
King Lear	*Goneril*	*Goneril*	**Action, imbalanced**
Macbeth	*Lady Macbeth*	--	**Action, Balanced**
Merchant of Venice	*Portia*	*Portia*	**Action & Romantic, Balanced**

Much Ado About Nothing	--	--	**Romantic, Balanced**
Othello	*Desdemona*	*Desdemona*	**Action & romantic, Balanced**
Romeo Juliet	--	--	**Romantic, Balanced,** *Nurse appears highly ranked in mentioning as she serves as a proxy to Juliet.*
Taming of the Shrew	*Katharina*	--	**Action & Romantic, Balanced**
Tempest	--	--	**Romantic, Balanced**
Twelfth Night	*Viola*	*Olivia, Viola*	**Romantic, Balanced**

Chapter 4

Time Series Analysis

This is a further step after creation of networks in our analysis. In time series analysis the list of five to ten important characters in the play are considered based on the analysis using Gephi. Based on the list, the respective edge pairs are extracted from the existing network/edge lists. Time series analysis is performed between scenes of the play considered.

Based on the important characters considered, the initial edge lists are filtered such that the resulting edge lists contain only the important characters. This extraction is done in both interaction and mentioning. These extracted edge lists have the ACT and SCENE division between them.

Later the extracted edge lists are considered and observed in the scene to scene fashion and as a result the roles of important characters are monitored from the start to the end of the play.

Time Series analysis is conducted on all the twelve comedy/tragedy plays we chose. Pictorial representation and description about the time series analysis for a play is as follows.

4.1 Hamlet

There are total of 5 ACT's the play.

4.1.1 Interaction

The list of important characters considered for the time series analysis for the play Hamlet – interaction edge list are

1) HORATIO

2) KINGCLAUDIUS

3) LORDPOLONIUS

4) HAMLET

5) QUEENGERTRUDE

Figure 6: *4.1 Hamlet Interaction- ACT I*

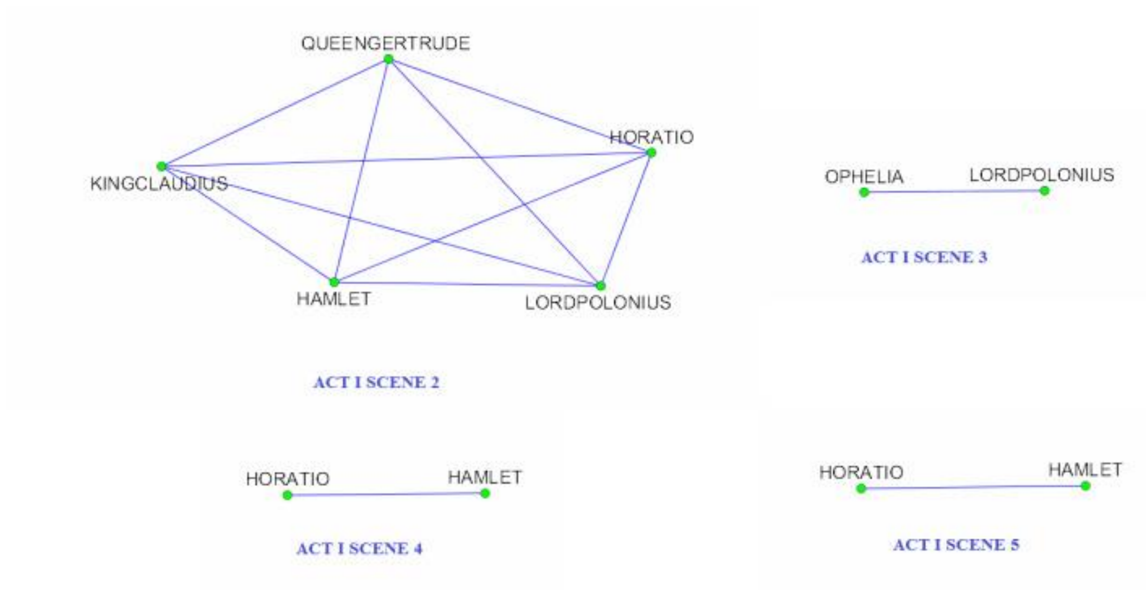


Figure 7: *4.2 Hamlet Interaction - ACT II*

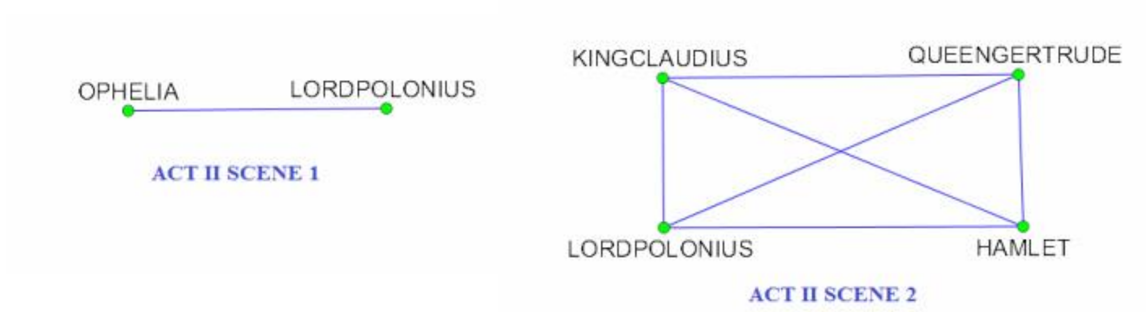


Figure 8: *4.3 Hamlet Interaction - ACT III*

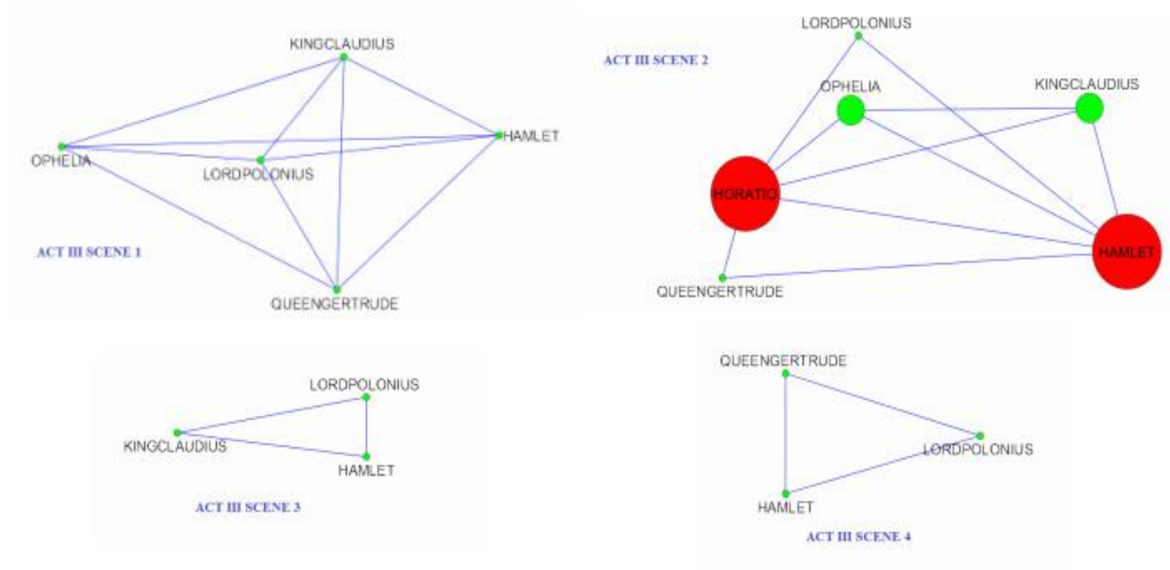


Figure 9: *4.4 Hamlet Interaction - ACT IV*

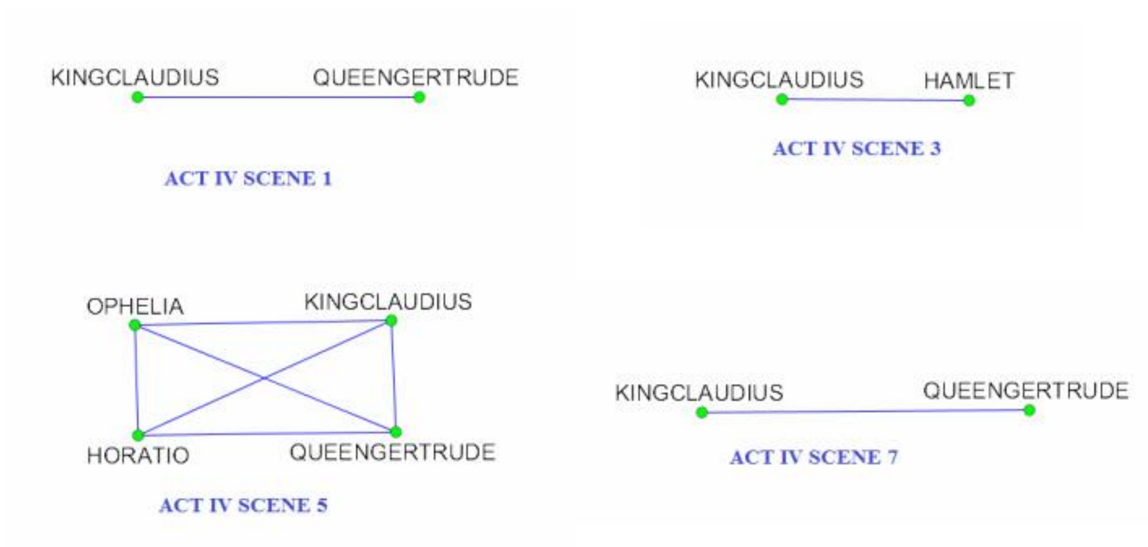
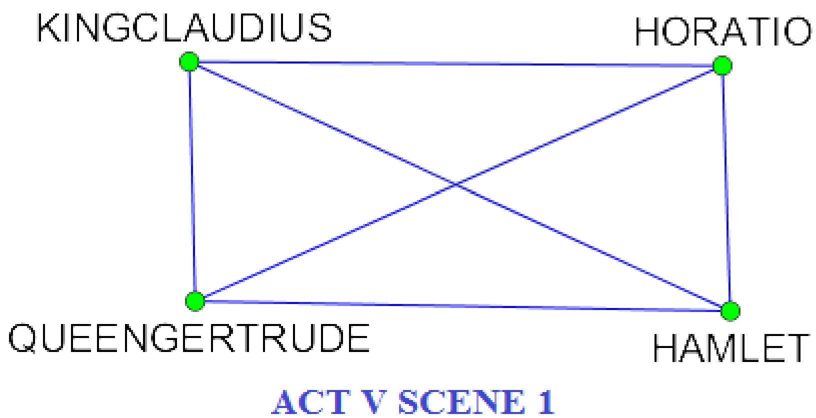


Figure 10: *4.5 Hamlet Interaction - ACT V*



Time series analysis is performed on *As You Like It* and *Hamlet*. From the above pictures for Hamlet- Interaction we can conclude that, each act has a gathering (closed figure) where important people talk. Also when we observe the above pictures, we can confirm that the important characters talk frequently.

For example, Ophelia, Lord Polonius, King Claudius, Hamlet and Queen Gertrude talk more when we observe each act.

Time series analysis is where we consider a set of important characters and monitor their influence on the play scene by scene. When interaction and mentioning models are considered we show multiple pairs of people appear frequently, where as in time series analysis we confine the list of people to the top five to 10 characters and monitor their role throughout the play.

For example, when the play "Hamlet" is considered we can see that King Claudius and Queen Gertrude appears in 7 scenes out of 15 which is significant. When we consider time-series, the analysis is confined to a maximum of 15 scenes for any play. When we consider entire play we only know how many lines they are interacting. But, not the number of times they are coming together.

For example, when we consider Romeo Juliet, we know that Romeo and Juliet talk much when they meet. But, they don't meet frequently. In fact, Juliet talks to nurse more when compared to Romeo.

The time series analysis for *As you like it* can be found in the appendix section.

Chapter 5

Software

5.1 Introduction:

For the user to extract the edge lists from the plays in a hassle free manner, we have developed a web tool which reads the plays and outputs the multiple types of edge lists with a single click. For the web tool developed, the plays can be read either from the file which is saved in html format in the local machine or directly from the website by providing the link. The different types of edge lists that can be extracted using this webpage are 1) Interaction with ACT and SCENE, the output consists of interaction edge lists separated with acts and scenes 2) Interaction, this consists of the combined interaction of pairs of characters appearing in the whole play. 3) Mentioning with ACT and SCENE, the output consists of mentioning edge lists which are separated with acts and scene's 4) Mentioning, the output consists of the mentioning counts of the pairs of characters combined throughout the play. The Mentioning edge lists along with the relationships can also be retrieved by reading a file containing the list of relationships into the additional textbox provided.

5.2 Screens

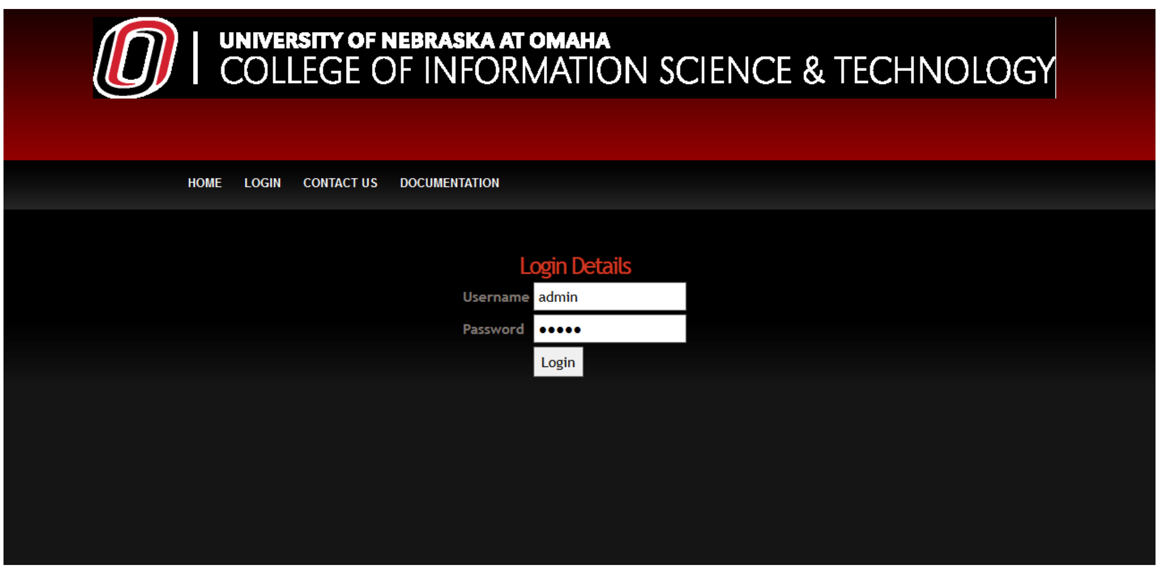
The screenshots of the software are below which explains the process step by step.

5.2.1 Home Page

Home page consists of three buttons Home, Login, Contact us and Documentation

5.2.2 Login Page

We can navigate to below screen by clicking the Login button.



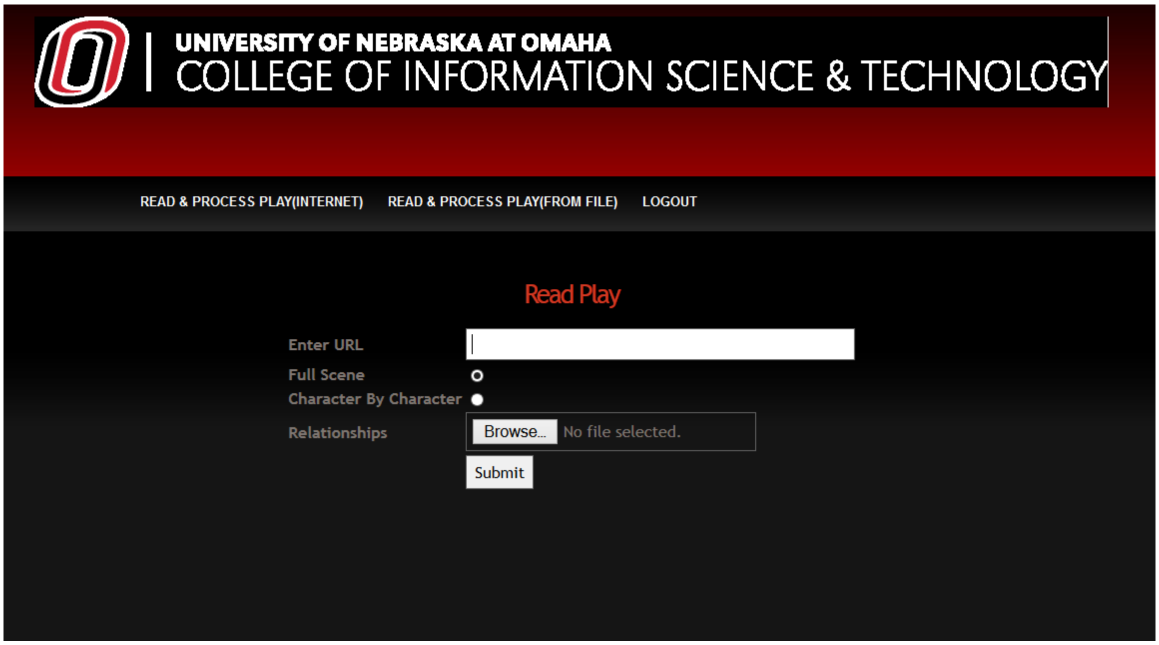
Here in this page, the users are supposed to provide login credentials to proceed further.

After the user is able to login successfully, they are directed to the page below.

In this page they see three links which says reading plays from web, reading plays from file and log out. By Clicking the read & Process play (Internet) or Read & Process play (From File) link. They will be navigated to the page where they actually can select the play either from web or local machine.

5.2.3 Read Plays from Internet

The below screen is reading the plays from internet



Here, in this screen we are able to enter the URL of the play we want to process and get the output. Then a radio button which is used to select between the Full Scene and Character by character edge lists of interaction and mentioning. Here the relationships field is not mandatory field.

The Relationships text field is used to read the file consisting of a list of relationships and pronouns which are included in the mentioning edge lists. On providing the link, selecting the radio button and clicking the submit button the control is directed to the page below.

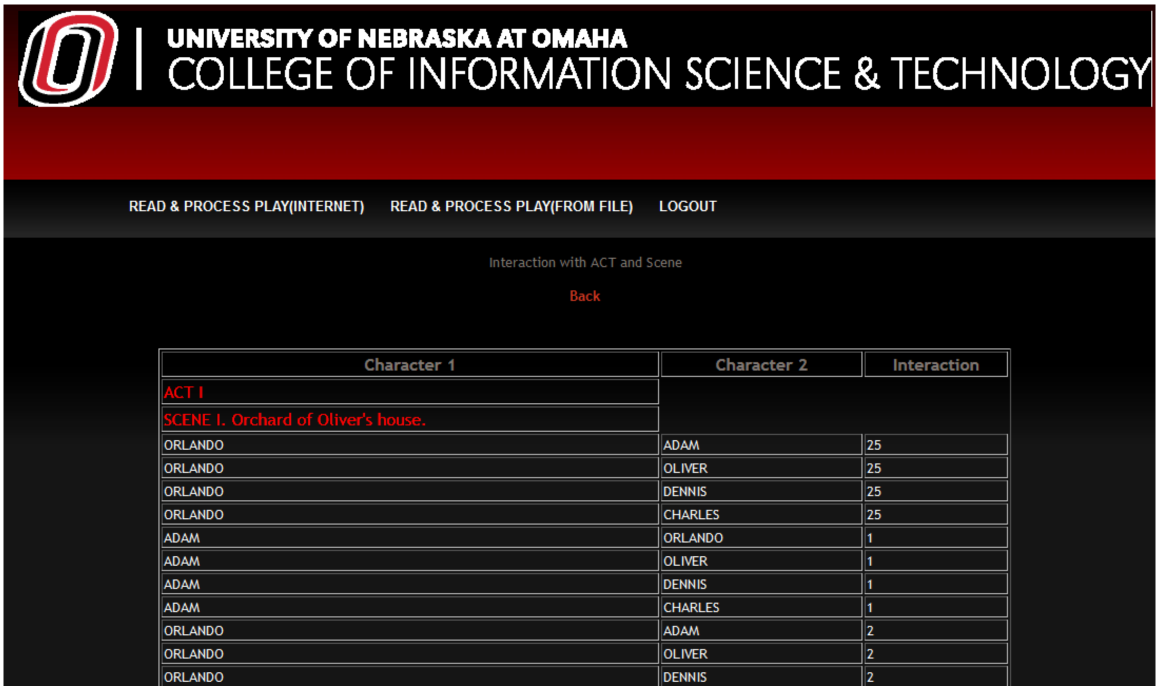
5.2.4 Results Page

Here the four buttons upon clicking give four different types of edge lists as explained earlier. Here the edge lists are displayed on the screen as well as saved in a text file in the background. For example if the play is *As You Like It* and an Interaction with Act and Scene is the type of edge network and full scene is selected then the saved text file name will be "asyoulikeit_Interaction with ACT and Scene_FullScene."

The output which is shown on the screen is in tabular form.

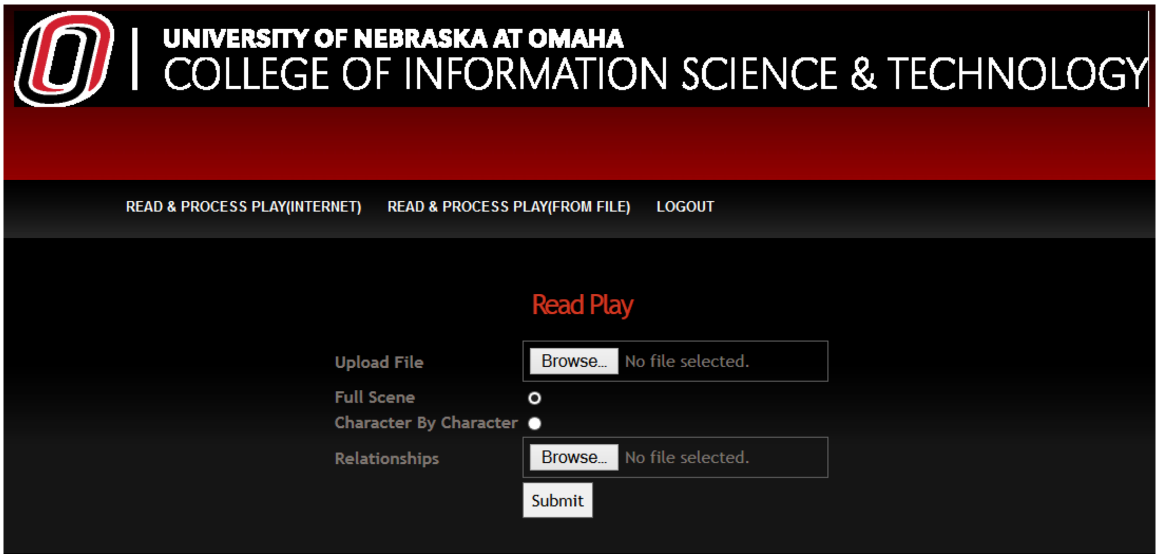
5.2.5 Results

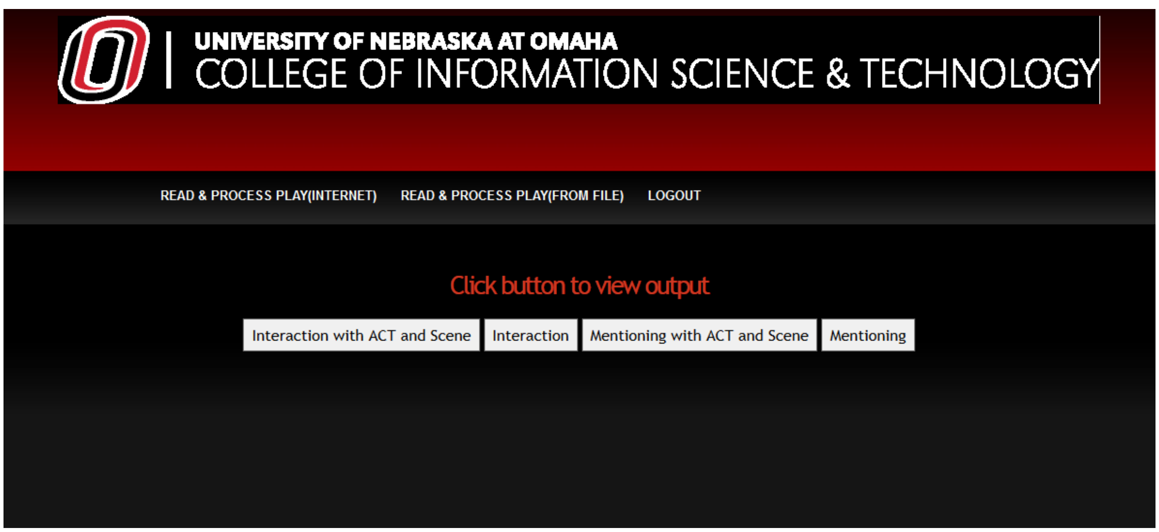
For example, when the link for the *As you like it* play is provided and the submit button is clicked, the control goes to the next page where you click the *Interaction with ACT and SCENE* button. An Image of the output is as follows.



5.2.6 Read Play from File

A screenshot of the *Read and Process play (from file)* follows. With two fields where we can read the html format of the play saved in the local machine and the relationships text file upon clicking the submit button will redirect to the results page.





5.2.7 Documentation

This page has all the information about the functionality of the webtool.

Chapter 6

Conclusion and Future Work

In this thesis we have analyzed Shakespeare plays in multiple network perspectives. Our question is, can we find the important characters of plays based on networks? And the answer is yes. Our different methods show different results based on the analysis we did. For example, interaction edge lists shows the social context [11] [12] where the women are less important and the messengers have high betweenness centrality, whereas the mentioning networks show the plot context. The important characters are mentioned often, socially peripheral people are mentioned less often, and female protagonist get high mentioning in romantic plots [16].

As we created multiple networks from single data sets, they provide more accurate information about what a social network is and have more scope to examine multiple characteristics. More interaction is not a necessary characteristic to calculate importance. Distribution of interaction and importance shows patterns of social relations.

Further, we created a new model called time-series analysis where the plays are broken down and the influence of important characters on the plot is studied. Therefore we are able to examine their behavior from the beginning towards the end of the play.

Multiple tools (namely Gephi and Cytoscape) are used to picture the networks both pictographically and in tabular forms by mapping the edge lists to multiple network metrics.

Finally, we found the plays have multiple underlying characteristics which cannot be found just by seeing them as a reader, we found them by plotting networks from plays [16].

In the future, we would like to research a new metric which can incorporate the pronouns/second names of the characters and also extend the research on different forms of data sets like movie scripts [13] and dynamic data sets which show change in the world/society with time such as twitter data and real life conversations. Hence it could be useful for legal findings and entertainment purposes.

Chapter 7

Appendix

Following are the tables of data in the Gephi Analysis

7.1 Interaction

Table 5: *7.1.1 As you Like it Interaction*

1	Id	Degree	RANK	Closeness Centrality	RANK	Betweenness Centrality	RANK	Eigenvector Centrality	RANK	AVERAGE RANK
2	ORLANDO	19	22	1.333333333	1	28.86825397	21	1	22	16.5
3	ROSALIND	18	21	1.476190476	3	23.09761905	20	0.828246525	21	16.25
4	CELIA	16	19	1.523809524	5	7.214285714	16	0.812360651	20	15
5	TOUCHSTONE	15	18	1.476190476	3	21.48095238	19	0.792126691	19	14.75
6	DUKEFREDERICK	10	17	1.666666667	6	12.83571429	18	0.706492943	16	14.25
7	OLIVER	8	10	1.80952381	11	6.325396825	15	0.628658725	15	12.75
8	ADAM	9	13	1.666666667	6	8.321428571	17	0.558579483	13	12.25
9	CHARLES	9	13	1.666666667	6	5.211111111	14	0.724069892	17	12.5
10	FirstLord	9	13	1.761904762	9	4.022222222	13	0.481718382	11	11.5
11	DUKESENIOR	7	9	1.80952381	11	2.505555556	10	0.408213773	8	9.5
12	AMIENS	8	10	1.80952381	11	2.505555556	10	0.408213773	8	9.75
13	SecondLord	4	6	2.19047619	19	0.666666667	8	0.237264673	6	9.75
14	CORIN	9	13	1.761904762	9	1.492857143	9	0.539960892	12	10.75
15	DENNIS	4	6	2.095238095	18	0	1	0.338952813	7	8
16	SILVIUS	8	10	2	15	3.733333333	12	0.435876407	10	11.75
17	JAQUES	17	20	1.380952381	2	63.71904762	22	0.761235599	18	15.5
18	LEBEAU	6	8	1.80952381	11	0	1	0.564033699	14	8.5
19	AUDREY	3	4	2.047619048	16	0	1	0.208184726	4	6.25
20	SIROLIVERMARTEXT	3	4	2.047619048	16	0	1	0.208184726	4	6.25
21	PHEBE	2	1	2.380952381	22	0	1	0.147071523	3	6.75
22	ALord	2	1	2.285714286	20	0	1	0.104238894	1	5.75
23	Forester	2	1	2.285714286	20	0	1	0.104238894	1	5.75

Table 6: *7.1.2 Hamlet Interaction*

1	Id	Degree	RANK	Eigenvect	RANK	Closeness	RANK	Betweenness Centrality	RANK	AVERAGE RANK
2	QUEENGERTRUDE	35	34	1	35	1.2	2	55.70576091	33	26
3	HAMLET	35	34	0.971954	32	1.228571	3	66.12114552	34	25.75
4	HORATIO	31	32	0.990116	34	1.142857	1	151.5255661	35	25.5
5	KINGCLAUDIUS	34	33	0.984218	33	1.228571	3	48.23909424	32	25.25
6	LORDPOLONIUS	24	31	0.783882	30	1.457143	6	26.96720779	31	24.5
7	All	23	29	0.917426	31	1.342857	5	20.66309524	28	23.25
8	OPHELIA	21	26	0.729667	29	1.514286	8	21.61804029	30	23.25
9	LAERTES	23	29	0.690658	27	1.457143	6	21.60201465	29	22.75
10	ROSENCRANTZ	22	28	0.714953	28	1.514286	8	18.91258741	27	22.75
11	GUILDENSTERN	21	26	0.677892	26	1.571429	10	7.868631369	25	21.75
12	FirstPlayer	19	25	0.668894	25	1.6	11	1.467532468	23	21
13	MARCELLUS	12	19	0.454232	19	1.685714	16	12.19393939	26	20
14	PRINCEFORTINBRAS	12	19	0.483529	20	1.714286	17	6.115384615	24	20
15	Prologue	13	21	0.64752	21	1.628571	12	0	1	13.75
16	PlayerKing	13	21	0.64752	21	1.628571	12	0	1	13.75
17	PlayerQueen	13	21	0.64752	21	1.628571	12	0	1	13.75
18	LUCIANUS	13	21	0.64752	21	1.628571	12	0	1	13.75
19	VOLTIMAND	7	11	0.362087	11	1.914286	28	0	1	12.75
20	OSRIC	9	16	0.433162	16	1.771429	18	0	1	12.75
21	Lord	9	16	0.433162	16	1.771429	18	0	1	12.75
22	FirstAmbassador	9	16	0.433162	16	1.771429	18	0	1	12.75
23	FirstClown	8	12	0.398326	12	1.8	22	0	1	11.75
24	SecondClown	8	12	0.398326	12	1.8	22	0	1	11.75
25	FirstPriest	8	12	0.398326	12	1.8	22	0	1	11.75
26	CORNELIUS	8	12	0.42577	15	1.771429	18	0	1	11.5
27	Gentleman	6	9	0.294293	9	1.885714	26	0	1	11.25
28	Danes	6	9	0.294293	9	1.885714	26	0	1	11.25
29	Messenger	3	4	0.167714	7	2.114286	33	0	1	11.25
30	Captain	3	4	0.135901	6	2.142857	34	0	1	11.25
31	Ghost	5	8	0.263474	8	1.857143	25	0	1	10.5
32	REYNALDO	2	1	0.094356	3	2.342857	35	0	1	10
33	BERNARDO	3	4	0.098017	4	2.057143	29	0	1	9.5
34	FRANCISCO	3	4	0.098017	4	2.057143	29	0	1	9.5
35	Servant	2	1	0.067246	1	2.085714	31	0	1	8.5
36	FirstSailor	2	1	0.067246	1	2.085714	31	0	1	8.5

Table 7: *7.1.3 Julius Caesar Interaction*

1	Id	Degree	RANK	Closeness Centrality	RANK	Betweenness Centrality	RANK	Eigenvector Centrality	RANK	AVERAGE RANK
2	BRUTUS	58	51	1.066666667	6	305.3718975	51	1	51	39.75
3	CASSIUS	41	49	1.311111111	8	121.3274531	49	0.847286307	50	39
4	ANTONY	41	49	1.266666667	7	172.7589827	50	0.825637845	49	38.75
5	Servant	24	47	1.533333333	9	35.06666667	48	0.631721054	48	38
6	CASCA	23	46	1.622222222	11	12.63134921	44	0.604890602	47	37
7	LUCILIUS	26	48	1.555555556	10	34.4508658	47	0.481189721	36	35.25
8	CINNA	19	40	1.644444444	13	11.19801587	43	0.581959487	44	35
9	LUCIUS	19	40	1.666666667	16	17.77777778	45	0.511802354	39	35
10	DECIUSBRUTUS	20	43	1.644444444	13	3.659126984	36	0.597665648	45	34.25
11	TREBONIUS	20	43	1.644444444	13	3.659126984	36	0.597665648	45	34.25
12	MESSALA	22	45	1.622222222	11	21.26753247	46	0.439573004	35	34.25
13	METELLUSCIMBER	19	40	1.666666667	16	2.725793651	35	0.574734534	42	33.25
14	Soothsayer	18	39	1.666666667	16	2.379761905	34	0.578213023	43	33
15	OCTAVIUS	15	36	1.777777778	23	10.68546176	42	0.300327937	30	32.75
16	PUBLIUS	16	38	1.711111111	19	0.4	27	0.52761968	40	31
17	TITINIUS	14	35	1.8	24	2	33	0.304115555	31	30.75
18	CAESAR	15	36	1.711111111	19	0.4	27	0.52761968	40	30.5
19	PORTIA	11	27	1.822222222	30	0.111111111	26	0.374960642	33	29
20	PINDARUS	11	27	1.844444444	35	1.4	32	0.252051927	22	29
21	CATO	10	24	1.822222222	30	1.278210678	31	0.288325543	29	28.5
22	ThirdCitizen	12	32	1.8	24	10.25	38	0.249583914	18	28
23	FirstSoldier	10	24	1.844444444	35	0.7254329	29	0.255740373	23	27.75
24	FirstCitizen	11	27	1.8	24	10.25	38	0.249583914	18	26.75
25	SecondCitizen	11	27	1.8	24	10.25	38	0.249583914	18	26.75
26	FourthCitizen	10	23	1.8	24	10.25	37	0.249583914	17	25.25
27	SecondSoldier	8	11	1.844444444	34	0.7254329	28	0.255740373	22	23.75
28	ARTEMIDORUS	13	32	1.733333333	21	0	1	0.504688565	36	22.5
29	POPILIUS	13	32	1.733333333	21	0	1	0.504688565	36	22.5
30	CALPURNIA	11	26	1.8	24	0	1	0.405727405	33	21
31	LIGARIUS	9	16	1.844444444	34	0	1	0.34244583	31	20.5
32	Poet	9	16	1.844444444	34	0	1	0.2721128	24	18.75
33	VARRO	9	16	1.844444444	34	0	1	0.2721128	24	18.75
34	GHOST	9	16	1.844444444	34	0	1	0.2721128	24	18.75
35	CLAUDIUS	9	16	1.844444444	34	0	1	0.2721128	24	18.75
36	Messenger	6	9	1.888888889	46	0	1	0.24005356	14	17.5
37	ThirdSoldier	6	9	1.911111111	47	0	1	0.191065515	9	16.5
38	CLITUS	8	11	1.866666667	42	0	1	0.23325158	10	16
39	DARDANIUS	8	11	1.866666667	42	0	1	0.23325158	10	16
40	VOLUMNIUS	8	11	1.866666667	42	0	1	0.23325158	10	16
41	STRATO	8	11	1.866666667	42	0	1	0.23325158	10	16
42	Citizens	9	16	1.822222222	30	0	1	0.245191928	15	15.5
43	SeveralCitizens	9	16	1.822222222	30	0	1	0.245191928	15	15.5
44	CICERO	3	2	2.244444444	49	0	1	0.122189611	8	15
45	CINNATHEPOET	4	3	2.711111111	50	0	1	0.064000883	6	15
46	LEPIDUS	2	1	2.222222222	48	0	1	0.069483916	7	14.25
47	FLAVIUS	4	3	1	1	0	1	0.006578504	1	1.5
48	FirstCommoner	4	3	1	1	0	1	0.006578504	1	1.5
49	MARULLUS	4	3	1	1	0	1	0.006578504	1	1.5
50	SecondCommoner	4	3	1	1	0	1	0.006578504	1	1.5

Table 8: *7.1.4 King Lear Interaction*

1	Id	Degree	RANK	Closeness Centrality	RANK	Betweenness Centrality	RANK	Eigenvector Centrality	RANK	AVERAGE RANK
2	GONERIL	29	26	1.12	1	26.19126984	25	1	26	19.5
3	GLOUCESTER	28	24	1.2	3	29.34285714	26	0.908785977	22	18.75
4	REGAN	28	24	1.16	2	21.41031746	24	0.975867525	25	18.75
5	KENT	27	23	1.24	4	13.05793651	21	0.928672948	23	17.75
6	EDGAR	23	20	1.36	8	13.92460317	23	0.804974275	19	17.5
7	KINGLEAR	25	21	1.24	4	13.05793651	21	0.928672948	23	17.25
8	EDMUND	25	21	1.28	6	11.09722222	19	0.889070635	21	16.75
9	ALBANY	22	19	1.28	6	9.511507937	18	0.888424523	20	15.75
10	OSWALD	19	16	1.36	8	11.86190476	20	0.763006337	16	15
11	Gentleman	19	16	1.4	11	5.61031746	15	0.794880659	18	15
12	CORDELIA	19	16	1.36	8	8.129365079	17	0.793887258	17	14.5
13	CORNWALL	18	15	1.44	12	6.728571429	16	0.682046685	15	14.5
14	Fool	15	14	1.56	13	1.276190476	13	0.617993761	14	13.5
15	Messenger	6	4	1.8	22	1.3	14	0.314571914	4	11
16	Doctor	5	2	2	25	0.5	12	0.257242326	2	10.25
17	Knight	6	4	1.84	23	0	1	0.350663835	8	9
18	Captain	10	9	1.6	14	0	1	0.587061704	12	9
19	Herald	10	9	1.6	14	0	1	0.587061704	12	9
20	LEAR	10	9	1.6	14	0	1	0.579229824	9	8.25
21	BURGUNDY	10	9	1.6	14	0	1	0.579229824	9	8.25
22	KINGOFFRANCE	10	9	1.6	14	0	1	0.579229824	9	8.25
23	FirstServant	7	6	1.76	19	0	1	0.345357265	5	7.75
24	SecondServant	7	6	1.76	19	0	1	0.345357265	5	7.75
25	ThirdServant	7	6	1.76	19	0	1	0.345357265	5	7.75
26	CURAN	5	2	1.88	24	0	1	0.291850724	3	7.5
27	OldMan	2	1	2.04	26	0	1	0.117421118	1	7.25

Table 9: *7.1.5 Macbeth Interaction*

1	Id	Degree	RANK	Closeness Centrality	RANK	Betweenness Centrality	RANK	Eigenvector Centrality	RANK	AVERAGE RANK
2	MACBETH	38	40	1.25	1	194.4716894	40	1	40	30.25
3	LENNOX	28	39	1.4	2	122.7729104	39	0.878891944	39	29.75
4	BANQUO	25	38	1.5	4	64.55699023	36	0.771100631	38	29
5	ROSS	24	37	1.475	3	100.3434538	38	0.759667227	37	28.75
6	LADYMACBETH	22	36	1.525	5	66.88200133	37	0.68133758	36	28.5
7	MALCOLM	21	35	1.6	6	36.71253746	34	0.614592283	35	27.5
8	MACDUFF	18	34	1.625	7	40.35301365	35	0.581738419	34	27.5
9	FirstMurderer	16	33	1.7	8	35.59010989	33	0.442876862	29	25.75
10	ANGUS	10	28	1.775	12	11.93455433	31	0.47023171	30	25.25
11	SecondWitch	14	31	1.75	9	3.777777778	25	0.561397945	31	24
12	ThirdWitch	14	31	1.75	9	3.777777778	25	0.561397945	31	24
13	FirstWitch	12	30	1.75	9	3.777777778	25	0.561397945	31	23.75
14	SIWARD	10	28	1.875	16	8.016433566	29	0.327190116	21	23.5
15	SecondMurderer	9	23	1.975	24	3.219444444	24	0.270168657	18	22.25
16	MENTEITH	7	17	2.1	29	7.496403596	28	0.238361473	14	22
17	Doctor	8	18	1.875	16	14.37777778	32	0.306639875	20	21.5
18	Messenger	8	18	1.925	19	8.497130647	30	0.250967127	15	20.5
19	DUNCAN	8	18	1.825	13	2.319047619	23	0.371801759	22	19
20	Servant	5	13	2.125	31	0.25	21	0.16246317	11	19
21	HECATE	9	23	1.925	19	0	1	0.420578371	23	16.5
22	FirstApparition	9	23	1.925	19	0	1	0.420578371	23	16.5
23	SecondApparition	9	23	1.925	19	0	1	0.420578371	23	16.5
24	ThirdApparition	9	23	1.925	19	0	1	0.420578371	23	16.5
25	SEYTON	4	7	2.075	28	0.833333333	22	0.132520762	8	16.25
26	Porter	8	18	1.85	14	0	1	0.433670407	27	15
27	DONALBAIN	8	18	1.85	14	0	1	0.433670407	27	15
28	ATTENDANT	6	15	2	25	0	1	0.261628163	16	14.25
29	BothMurderers	6	15	2	25	0	1	0.261628163	16	14.25
30	Soldiers	4	7	2.425	39	0	1	0.134816268	10	14.25
31	Lords	5	13	1.9	18	0	1	0.285932879	19	12.75
32	Sergeant	4	7	2.1	29	0	1	0.198714465	13	12.5
33	ThirdMurderer	3	5	2.325	37	0	1	0.113598663	4	11.75
34	LADYMACDUFF	4	7	2.25	33	0	1	0.122061099	6	11.75
35	Son	4	7	2.25	33	0	1	0.122061099	6	11.75
36	YOUNGSIWARD	4	7	2.05	27	0	1	0.191881105	12	11.75
37	CAITHNESS	3	5	2.275	35	0	1	0.120313784	5	11.5
38	Gentlewoman	2	2	2.45	40	0	1	0.075783923	2	11.25
39	FLEANCE	2	2	2.175	32	0	1	0.133945121	9	11
40	OldMan	2	2	2.3	36	0	1	0.101904096	3	10.5
41	Lord	1	1	2.375	38	0	1	0.06620211	1	10.25

Table 10: *7.1.6 Merchant Interaction*

1	Id	Degree	RANK	Closeness Centrality	RANK	Betweenness Centrality	RANK	Eigenvector Centrality	RANK	AVERAGE RANK
2	PORTIA	27	22	1.095238095	1	37.96230159	22	1	22	16.75
3	GRATIANO	26	21	1.142857143	2	19.29563492	20	0.992675465	19	15.5
4	NERISSA	23	19	1.142857143	2	17.96230159	19	0.993572746	21	15.25
5	BASSANIO	23	19	1.142857143	2	19.29563492	20	0.992675465	19	15
6	LORENZO	21	17	1.333333333	6	5.763095238	17	0.842714939	17	14.25
7	ANTONIO	21	17	1.285714286	5	4.602777778	16	0.918348416	18	14
8	LAUNCELOT	17	15	1.380952381	7	8.623015873	18	0.777275089	15	13.75
9	JESSICA	19	16	1.380952381	7	4.41468254	15	0.806220203	16	13.5
10	SALARINO	15	14	1.476190476	9	0.813492063	11	0.74070722	14	12
11	Servant	11	10	1.571428571	12	2.5	14	0.585566315	9	11.25
12	SHYLOCK	11	10	1.476190476	9	1.202380952	12	0.722575376	13	11
13	SALERIO	11	10	1.476190476	9	1.43968254	13	0.710175795	12	11
14	SALANIO	11	10	1.571428571	12	0.125	10	0.627541061	11	10.75
15	STEPHANO	8	7	1.619047619	14	0	1	0.58587137	10	8
16	ALL	7	6	1.666666667	17	0	1	0.507123906	6	7.5
17	DUKE	8	7	1.619047619	14	0	1	0.550416986	7	7.25
18	Clerk	8	7	1.619047619	14	0	1	0.550416986	7	7.25
19	BALTHASAR	4	3	1.904761905	18	0	1	0.29143287	5	6.75
20	GOBBO	4	3	1.952380952	19	0	1	0.24114568	3	6.5
21	LEONARDO	4	3	1.952380952	19	0	1	0.24114568	3	6.5
22	MOROCCO	2	1	2.047619048	22	0	1	0.08019016	1	6.25
23	ARRAGON	3	2	1.952380952	19	0	1	0.206666928	2	6

Table 11:*7.1.7 Much Ado About Nothing Interaction*

1	Id	Degree	RANK	Closeness Centrality	RANK	Betweenness Centrality	RANK	Eigenvector Centrality	RANK	AVERAGE RANK
2	LEONATO	24	22	1.272727273	2	16.25045788	22	1	23	17.25
3	DONPEDRO	22	19	1.318181818	3	13.32554945	20	0.954316662	21	15.75
4	CLAUDIO	22	19	1.318181818	3	13.32554945	20	0.954316662	20	15.5
5	BORACHIO	20	16	1.181818182	1	62.64010989	23	0.967257289	22	15.5
6	BENEDICK	25	23	1.363636364	5	3.325549451	14	0.942755173	18	15
7	BEATRICE	23	21	1.363636364	5	3.325549451	14	0.942755173	18	14.5
8	DONJOHN	21	17	1.363636364	5	8.602289377	19	0.934025442	17	14.5
9	Messenger	14	11	1.5	11	6.782051282	18	0.713517217	11	12.75
10	HERO	21	17	1.409090909	8	1.282692308	10	0.90971384	16	12.75
11	BALTHASAR	16	14	1.454545455	10	1.642857143	12	0.841571557	14	12.5
12	ANTONIO	16	14	1.409090909	8	1.282692308	10	0.90971384	15	11.75
13	VERGES	11	10	1.636363636	14	5.203479853	16	0.358000509	7	11.75
14	DOGBERRY	10	8	1.636363636	14	5.203479853	16	0.358000509	7	11.25
15	CONRADE	9	6	1.727272727	16	2.807692308	13	0.299919343	6	10.25
16	MARGARET	15	13	1.5	11	0	1	0.808530224	12	9.25
17	FRIARFRANCIS	10	8	1.772727273	17	0	1	0.656272391	10	9
18	URSULA	14	11	1.5	11	0	1	0.808530224	12	8.75
19	Boy	6	2	1.954545455	22	0	1	0.447471095	9	8.5
20	SecondWatchman	9	6	1.818181818	18	0	1	0.231057959	2	6.75
21	FirstWatchman	8	5	1.818181818	18	0	1	0.231057959	2	6.5
22	Lord	2	1	2.227272727	23	0	1	0.15156515	1	6.5
23	Watchman	7	3	1.818181818	18	0	1	0.231057959	2	6
24	Sexton	7	3	1.818181818	18	0	1	0.231057959	2	6

Table 12: *7.1.8 Othello Interaction*

1	Id	Degree	RANK	Closeness Centrality	RANK	Betweenness Centrality	RANK	Eigenvector Centrality	RANK	AVERAGE RANK
2	IAGO	36	25	1	1	48.35952381	25	1	25	19
3	OTHELLO	31	23	1.04	2	37.27619048	24	0.986328317	24	18.25
4	DESDEMONA	32	24	1.08	3	26.27619048	23	0.977045362	23	18.25
5	RODERIGO	27	21	1.12	4	22.20952381	22	0.953215494	22	17.25
6	CASSIO	27	21	1.24	5	15.10952381	21	0.833814014	21	17
7	EMILIA	25	20	1.32	6	9.20952381	20	0.764150557	20	16.5
8	MONTANO	19	19	1.44	7	2.142857143	19	0.695381094	19	16
9	LODOVICO	13	15	1.6	15	0.25	13	0.535107686	12	13.75
10	GRATIANO	13	15	1.6	15	0.25	13	0.535107686	12	13.75
11	BIANCA	11	9	1.64	22	0.5	15	0.481428852	9	13.75
12	BRABANTIO	15	17	1.56	8	0.833333333	17	0.534118986	10	13
13	FirstOfficer	15	17	1.56	8	0.833333333	17	0.534118986	10	13
14	Clown	9	3	1.72	23	0.75	16	0.364053901	3	11.25
15	FirstGentleman	11	9	1.56	8	0	1	0.593450999	14	8
16	SecondGentleman	11	9	1.56	8	0	1	0.593450999	14	8
17	ThirdGentleman	11	9	1.56	8	0	1	0.593450999	14	8
18	FourthGentleman	11	9	1.56	8	0	1	0.593450999	14	8
19	SecondGentlemen	11	9	1.56	8	0	1	0.593450999	14	8
20	FirstMusician	4	2	1.84	24	0	1	0.20544915	2	7.25
21	Gentleman	2	1	1.92	25	0	1	0.137769339	1	7
22	DUKEOFVENICE	10	4	1.6	15	0	1	0.480114748	4	6
23	FirstSenator	10	4	1.6	15	0	1	0.480114748	4	6
24	SecondSenator	10	4	1.6	15	0	1	0.480114748	4	6
25	Sailor	10	4	1.6	15	0	1	0.480114748	4	6
26	Messenger	10	4	1.6	15	0	1	0.480114748	4	6

Table 13: *7.1.9 Romeo Juliet Interaction*

1	Id	Degree	RANK	Closeness Centrality	RANK	Betweenness Centrality	RANK	Eigenvector Centrality	RANK	AVERAGE RANK
2	ROMEO	38	33	1.151515152	3	67.39920635	34	0.959427357	32	25.5
3	CAPULET	40	34	1.121212121	1	56.38095238	33	0.997742996	33	25.25
4	LADYCAPULET	37	32	1.121212121	1	54.52142857	32	1	34	24.75
5	Nurse	28	31	1.424242424	8	21.80555556	30	0.654045398	24	23.25
6	FRIARLAURENCE	25	28	1.424242424	8	43.75079365	31	0.682320353	26	23.25
7	BENVOLIO	25	28	1.393939394	4	19.06031746	29	0.762057899	29	22.5
8	PRINCE	25	28	1.393939394	4	13.26587302	26	0.779012114	30	22
9	JULIET	24	27	1.424242424	8	11.93809524	25	0.753089533	28	22
10	MONTAGUE	23	25	1.393939394	4	13.26587302	26	0.779012114	30	21.25
11	TYBALT	21	24	1.484848485	11	9.292063492	24	0.679925328	25	21
12	PARIS	23	25	1.393939394	4	16.27301587	28	0.750218804	27	21
13	FirstCitizen	16	23	1.666666667	18	0.682539683	20	0.501381967	18	19.75
14	BALTHASAR	13	20	1.606060606	12	7.6	23	0.540526753	23	19.5
15	Servant	14	22	1.696969697	19	0.674603175	19	0.472603987	12	18
16	MERCUTIO	12	13	1.727272727	24	2.491269841	21	0.412090451	9	16.75
17	PETER	13	20	1.606060606	12	7.376190476	22	0.47435639	13	16.75
18	FirstServant	12	13	1.727272727	24	0.111111111	17	0.429687287	10	16
19	SecondServant	12	13	1.727272727	24	0.111111111	17	0.429687287	10	16
20	PAGE	12	13	1.636363636	14	0	1	0.534965386	19	11.75
21	FirstWatchman	12	13	1.636363636	14	0	1	0.534965386	19	11.75
22	SecondWatchman	12	13	1.636363636	14	0	1	0.534965386	19	11.75
23	ThirdWatchman	12	13	1.636363636	14	0	1	0.534965386	19	11.75
24	SAMPSON	11	9	1.696969697	19	0	1	0.477654282	14	10.75
25	GREGORY	11	9	1.696969697	19	0	1	0.477654282	14	10.75
26	ABRAHAM	11	9	1.696969697	19	0	1	0.477654282	14	10.75
27	LADYMONTAGUE	11	9	1.696969697	19	0	1	0.477654282	14	10.75
28	SecondCapulet	9	4	1.757575758	27	0	1	0.372499559	8	10
29	NURSE	5	3	1.878787879	32	0	1	0.198688997	3	9.75
30	Apothecary	2	2	2.090909091	33	0	1	0.090814221	2	9.5
31	FRIARJOHN	1	1	2.393939394	34	0	1	0.041508212	1	9.25
32	FirstMusician	9	4	1.757575758	27	0	1	0.340412576	4	9
33	SecondMusician	9	4	1.757575758	27	0	1	0.340412576	4	9
34	Musician	9	4	1.757575758	27	0	1	0.340412576	4	9
35	ThirdMusician	9	4	1.757575758	27	0	1	0.340412576	4	9

Table 14: *7.1.10 Taming of the Shrew Interaction*

1	Id	Degree	RANK	Closeness Centrality	RANK	Betweenness Centrality	RANK	Eigenvector Centrality	RANK	AVERAGE RANK
2	PETRUCHIO	30	33	1.527777778	7	36.25383368	31	0.873608221	32	25.75
3	KATHARINA	33	34	1.305555556	1	70.94238404	32	1	34	25.25
4	FirstServant	24	25	1.333333333	2	110.5535202	33	0.918063012	33	23.25
5	GRUMIO	25	29	1.611111111	12	23.52502415	28	0.773151322	24	23.25
6	HORTENSIO	24	25	1.472222222	4	34.91486125	30	0.855284092	31	22.5
7	LUCENTIO	27	30	1.5	5	17.17794959	25	0.854994991	29	22.25
8	BIANCA	27	30	1.5	5	17.17794959	25	0.854994991	29	22.25
9	TRANIO	28	32	1.527777778	7	8.675717374	21	0.835760487	25	21.25
10	VINCENTIO	16	22	1.972222222	22	0.1	20	0.574459114	21	21.25
11	SLY	22	23	1.388888889	3	128.976702	34	0.751866867	23	20.75
12	GREMIO	24	25	1.527777778	7	8.675717374	21	0.835760487	25	19.5
13	BIONDELLO	24	25	1.527777778	7	8.675717374	21	0.835760487	25	19.5
14	BAPTISTA	23	24	1.527777778	7	8.675717374	21	0.835760487	25	19.25
15	Lord	13	20	1.805555556	14	28.56923077	29	0.294269763	12	18.75
16	Servant	10	17	1.888888889	15	19.71669664	27	0.291965959	11	17.5
17	Pedant	13	20	2	26	0	1	0.541633278	19	16.5
18	Widow	10	17	2	26	0	1	0.541633278	19	15.75
19	HORTENSIA	11	19	1.694444444	13	0	1	0.603756068	22	13.75
20	CURTIS	9	11	1.916666667	16	0	1	0.352687473	13	10.25
21	NATHANIEL	9	11	1.916666667	16	0	1	0.352687473	13	10.25
22	PHILIP	9	11	1.916666667	16	0	1	0.352687473	13	10.25
23	JOSEPH	9	11	1.916666667	16	0	1	0.352687473	13	10.25
24	NICHOLAS	9	11	1.916666667	16	0	1	0.352687473	13	10.25
25	PETER	9	11	1.916666667	16	0	1	0.352687473	13	10.25
26	Haberdasher	5	1	2.138888889	28	0	1	0.244973244	9	9.75
27	Tailor	5	1	2.138888889	28	0	1	0.244973244	9	9.75
28	Hostess	7	3	2.194444444	30	0	1	0.135572115	1	8.75
29	FirstHuntsman	7	3	2.194444444	30	0	1	0.135572115	1	8.75
30	SecondHuntsman	7	3	2.194444444	30	0	1	0.135572115	1	8.75
31	Players	7	3	2.194444444	30	0	1	0.135572115	1	8.75
32	APlayer	7	3	2.194444444	30	0	1	0.135572115	1	8.75
33	SecondServant	7	3	1.972222222	22	0	1	0.225663573	6	8
34	ThirdServant	7	3	1.972222222	22	0	1	0.225663573	6	8
35	Messenger	7	3	1.972222222	22	0	1	0.225663573	6	8

Table 15: *7.1.11 Tempest Interaction*

1	Id	Degree	RANK	Closeness Centrality	RANK	Betweenness Centrality	RANK	Eigenvector Centrality	RANK	AVERAGE RANK
2	PROSPERO	18	18	1.166666667	1	36	18	1	19	14
3	ARIEL	22	19	1.166666667	1	36	18	1	18	14
4	SEBASTIAN	14	17	1.444444444	3	9	14	0.595244028	6	10
5	ANTONIO	12	13	1.444444444	3	9	14	0.595244028	6	9
6	GONZALO	12	13	1.444444444	3	9	14	0.595244028	6	9
7	MIRANDA	13	15	1.666666667	9	0	1	0.680683215	11	9
8	CALIBAN	13	15	1.666666667	9	0	1	0.680683215	11	9
9	ALONSO	11	9	1.444444444	3	9	14	0.595244028	6	8
10	FERDINAND	11	9	1.666666667	9	0	1	0.680683215	11	7.5
11	TRINCULO	11	9	1.666666667	9	0	1	0.680683215	11	7.5
12	STEPHANO	11	9	1.666666667	9	0	1	0.680683215	10	7.25
13	IRIS	9	4	1.666666667	9	0	1	0.680683215	11	6.25
14	CERES	9	4	1.666666667	9	0	1	0.680683215	11	6.25
15	JUNO	9	4	1.666666667	9	0	1	0.680683215	11	6.25
16	Master	6	1	2.111111111	17	0	1	0.316719774	1	5
17	Boatswain	6	1	2.111111111	17	0	1	0.316719774	1	5
18	Mariners	6	1	2.111111111	17	0	1	0.316719774	1	5
19	ADRIAN	10	8	1.611111111	7	0	1	0.503013757	4	5
20	FRANCISCO	9	4	1.611111111	7	0	1	0.503013757	4	4

Table 16: *7.1.12 Twelfth Night Interaction*

1	Id	Degree	RANK	Closeness Centrality	RANK	Betweenness Centrality	RANK	Eigenvector Centrality	RANK	AVERAGE RANK
2	VIOLA	17	14	1.0625	1	41.85714286	17	1	17	12.25
3	SIRTOBYBELCH	18	15	1.25	2	2.752380952	13	0.979513431	14	11
4	SIRANDREW	18	15	1.25	2	2.752380952	13	0.979513431	14	11
5	Clown	13	10	1.4375	12	8.228571429	16	0.630387021	6	11
6	MARIA	18	15	1.3125	5	0.80952381	10	0.942520041	12	10.5
7	OLIVIA	15	12	1.25	2	2.752380952	13	0.979513431	14	10.25
8	MALVOLIO	15	12	1.3125	5	0.80952381	10	0.942520041	12	9.75
9	ANTONIO	12	9	1.3125	5	1.942857143	12	0.924229168	11	9.25
10	SEBASTIAN	6	4	1.8125	15	0.142857143	7	0.42473023	5	7.75
11	DUKEORSINO	6	4	1.75	13	0.476190476	8	0.190484562	3	7
12	CURIO	5	3	1.75	13	0.476190476	8	0.190484562	3	6.75
13	FABIAN	14	11	1.375	8	0	1	0.887235778	7	6.75
14	Servant	10	6	1.375	8	0	1	0.887235778	7	5.5
15	FirstOfficer	10	6	1.375	8	0	1	0.887235778	7	5.5
16	SecondOfficer	10	6	1.375	8	0	1	0.887235778	7	5.5
17	VALENTINE	4	2	1.875	16	0	1	0.135200299	2	5.25
18	Captain	1	1	2	17	0	1	0.095477275	1	5

7.2 Mentioning

Table 17: *7.2.1 As you Like it Mentioning*

1	Id	In-Degree	RANK	Out-Degree	RANK	PageRank	RANK	AVERAGE RANK
2	ORLANDO	4	21	8	22	0.12720494	22	21.66666667
3	ROSALIND	5	22	7	21	0.11077918	21	21.33333333
4	TOUCHSTONE	3	15	4	19	0.07305959	20	18
5	OLIVER	2	10	4	19	0.07129308	19	16
6	DUKESENIOR	3	15	2	12	0.07013138	18	15
7	JAQUES	3	15	2	12	0.06438651	17	14.66666667
8	CELIA	3	15	2	12	0.05185573	16	14.33333333
9	CORIN	3	15	1	6	0.05166824	15	12
10	SILVIUS	2	10	2	12	0.03938195	13	11.66666667
11	PHEBE	3	15	1	6	0.03924689	12	11
12	WILLIAM	2	10	2	12	0.03346204	10	10.66666667
13	DUKEFREDERICK	1	4	2	12	0.04108453	14	10
14	AUDREY	2	10	1	6	0.03346204	10	8.666666667
15	HYMEN	0	1	2	12	0.02919709	8	7
16	LEBEAU	2	10	0	1	0.02926935	9	6.666666667
17	CHARLES	1	4	1	6	0.01894255	3	4.333333333
18	AMIENS	0	1	1	6	0.02051503	6	4.333333333
19	Forester	1	4	0	1	0.02051503	6	3.666666667
20	SIROLIVERMARTEXT	1	4	0	1	0.01923953	5	3.333333333
21	FirstLord	0	1	1	6	0.01873507	2	3
22	DENNIS	1	4	0	1	0.01894255	3	2.666666667
23	ADAM	1	4	0	1	0.01762775	1	2

Table 18: *7.2.2 Hamlet Mentioning*

1	Id	In-Degree	RANK	Out-Degree	RANK	PageRank	RANK	AVERAGE RANK
2	HAMLET	10	24	11	24	0.153633	24	24
3	KINGCLAUDIUS	6	23	7	23	0.109631	23	23
4	HORATIO	4	19	6	20	0.082009	22	20.33333333
5	LAERTES	5	21	3	18	0.061246	19	19.33333333
6	BERNARDO	3	18	3	18	0.037292	15	17
7	MARCELLUS	4	19	2	15	0.047158	17	17
8	OSRIC	2	15	2	15	0.043283	16	15.33333333
9	QUEENGERTRUDE	1	3	6	20	0.065603	20	14.33333333
10	LORDPOLONIUS	1	3	6	20	0.065603	20	14.33333333
11	OPHELIA	5	21	0	1	0.051902	18	13.33333333
12	ROSENCRANTZ	2	15	1	9	0.024616	11	11.66666667
13	FRANCISCO	1	3	2	15	0.026588	13	10.33333333
14	All	2	15	0	1	0.024616	11	9
15	PRINCEFORTINBRAS	1	3	1	9	0.033005	14	8.666666667
16	FirstSailor	0	1	1	9	0.015965	9	6.333333333
17	Ghost	1	3	1	9	0.015324	6	6
18	FirstClown	1	3	1	9	0.015324	6	6
19	Captain	1	3	0	1	0.020017	10	4.666666667
20	Lord	0	1	1	9	0.015324	6	5.333333333
21	CORNELIUS	1	3	0	1	0.015317	3	2.333333333
22	VOLTIMAND	1	3	0	1	0.015317	3	2.333333333
23	Danes	1	3	0	1	0.015317	3	2.333333333
24	REYNALDO	1	3	0	1	0.015292	1	1.666666667
25	GUILDENSTERN	1	3	0	1	0.015292	1	1.666666667

Table 19: *7.2.3 Julius Caesar Mentioning*

1	Id	In-Degree	RANK	Out-Degree	RANK	PageRank	RANK	AVERAGE RANK
2	BRUTUS	16	40	22	40	0.1464304	40	40
3	CAESAR	14	39	12	38	0.080131	39	38.66666667
4	ANTONY	13	38	9	37	0.0674649	38	37.66666667
5	CASSIUS	8	37	13	39	0.0615478	37	37.66666667
6	OCTAVIUS	5	35	6	36	0.0337353	36	35.66666667
7	MESSALA	5	35	4	33	0.0328836	35	34.33333333
8	CASCA	4	32	5	35	0.0256009	31	32.66666667
9	FirstCitizen	4	32	2	19	0.0262679	32	27.66666667
10	DECIUSBRUTUS	4	32	2	19	0.0253219	30	27
11	ARTEMIDORUS	2	15	3	29	0.0271795	34	26
12	FourthCitizen	2	15	4	33	0.0210887	29	25.66666667
13	LUCILIUS	2	15	3	29	0.0262793	33	25.66666667
14	STRATO	3	29	2	19	0.0205871	27	25
15	PINDARUS	3	29	2	19	0.0200151	25	24.33333333
16	ThirdCitizen	2	15	3	29	0.0210862	28	24
17	CINNA	3	29	2	19	0.0163196	21	23
18	TITINIUS	2	15	3	29	0.0200151	25	23
19	PORTIA	2	15	2	19	0.0172097	23	19
20	Servant	2	15	2	19	0.0166775	22	18.66666667
21	FirstSoldier	2	15	1	5	0.0199323	24	14.66666667
22	Soothsayer	1	4	2	19	0.0127343	16	13
23	METELLUSCIMBER	1	4	2	19	0.0125677	14	12.33333333
24	CLITUS	2	15	1	5	0.0139265	17	12.33333333
25	DARDANIUS	2	15	1	5	0.0139265	17	12.33333333
26	SecondCitizen	1	4	2	19	0.0121089	13	12
27	TREBONIUS	2	15	1	5	0.0125677	14	11.33333333
28	VARRO	2	15	0	1	0.0139265	17	11
29	LEPIDUS	2	15	1	5	0.0117341	12	10.66666667
30	CICERO	2	15	1	5	0.0116256	11	10.33333333
31	LUCIUS	1	4	1	5	0.0139265	17	8.666666667
32	CALPURNIA	1	4	1	5	0.0080923	7	5.333333333
33	PUBLIUS	1	4	1	5	0.0080923	7	5.333333333
34	SecondSoldier	0	1	1	5	0.0091936	10	5.333333333
35	FLAVIUS	0	1	1	5	0.0080923	7	4.333333333
36	LIGARIUS	1	4	1	5	0.0079888	2	3.666666667
37	CATO	1	4	1	5	0.0079888	2	3.666666667
38	GHOST	0	1	1	5	0.0079888	2	2.666666667
39	Poet	1	4	0	1	0.0079888	2	2.333333333
40	VOLUMNIUS	1	4	0	1	0.0079888	2	2.333333333
41	POPILIUS	1	4	0	1	0.0072653	1	2

Table 20: *7.2.4 King Lear Mentioning*

1	Id	In-Degree	RANK	Out-Degree	RANK	PageRank	RANK	AVERAGE RANK
2	KINGLEAR	7	21	12	21	0.1314761	21	21
3	EDMUND	6	20	5	18	0.1013058	20	19.33333333
4	GONERIL	5	17	5	18	0.0836326	19	18
5	GLOUCESTER	4	16	4	16	0.0677701	18	16.66666667
6	KENT	5	17	4	16	0.0672568	17	16.66666667
7	ALBANY	5	17	2	10	0.062346	15	14
8	REGAN	3	13	3	12	0.0630686	16	13.66666667
9	EDGAR	3	13	3	12	0.0458675	12	12.33333333
10	Fool	3	13	3	12	0.0341606	10	11.66666667
11	CORDELIA	1	2	5	18	0.054458	14	11.33333333
12	KINGOFFRANCE	2	8	3	12	0.0445936	11	10.33333333
13	CORNWALL	2	8	2	10	0.0491589	13	10.33333333
14	Captain	2	8	1	4	0.0341454	9	7
15	OSWALD	2	8	1	4	0.0257731	8	6.666666667
16	BURGUNDY	2	8	1	4	0.0252194	6	6
17	LEAR	1	2	1	4	0.0253332	7	4.333333333
18	SecondServant	1	2	0	1	0.0175826	5	2.666666667
19	Gentleman	1	2	1	4	0.0166771	1	2.333333333
20	OldMan	1	2	0	1	0.0167371	4	2.333333333
21	Herald	0	1	1	4	0.0167188	2	2.333333333
22	CURAN	1	2	0	1	0.0167188	2	1.666666667

Table 21:*7.2.5 Macbeth Mentioning*

1	Id	In-Degree	RANK	Out-Degree	RANK	PageRank	RANK	AVERAGE RANK
2	MACBETH	8	26	11	26	0.248284	26	26
3	MALCOLM	3	23	5	25	0.084746	25	24.33333333
4	MACDUFF	3	23	3	23	0.073738	24	23.33333333
5	FirstWitch	2	21	3	23	0.052157	22	22
6	BANQUO	3	23	1	12	0.053577	23	19.33333333
7	SecondWitch	2	21	1	12	0.039593	20	17.66666667
8	ROSS	1	6	2	20	0.045852	21	15.66666667
9	ThirdWitch	1	6	2	20	0.039236	19	15
10	DUNCAN	0	1	2	20	0.029364	18	13
11	HECATE	1	6	1	12	0.027853	17	11.66666667
12	Porter	0	1	1	12	0.018085	14	9
13	ANGUS	1	6	0	1	0.018552	16	7.666666667
14	SIWARD	1	6	0	1	0.018085	14	7
15	DONALBAIN	1	6	1	12	0.017558	1	6.333333333
16	FirstApparition	0	1	1	12	0.01798	3	5.333333333
17	SecondApparition	0	1	1	12	0.01798	3	5.333333333
18	ThirdApparition	0	1	1	12	0.01798	3	5.333333333
19	LADYMACBETH	1	6	0	1	0.01798	3	3.333333333
20	FirstMurderer	1	6	0	1	0.01798	3	3.333333333
21	BothMurderers	1	6	0	1	0.01798	3	3.333333333
22	LENNOX	1	6	0	1	0.01798	3	3.333333333
23	Servant	1	6	0	1	0.01798	3	3.333333333
24	SEYTON	1	6	0	1	0.01798	3	3.333333333
25	Doctor	1	6	0	1	0.01798	3	3.333333333
26	YOUNGSIWARD	1	6	0	1	0.01798	3	3.333333333
27	Sergeant	1	6	0	1	0.017558	1	2.666666667

Table 22:*7.2.6 Merchant Mentioning*

1	Id	In-Degree	RANK	Out-Degree	RANK	PageRank	RANK	AVERAGE RANK
2	PORTIA	8	23	9	23	0.140586	23	23
3	BASSANIO	7	21	7	21	0.103337	22	21.33333333
4	LORENZO	7	21	7	21	0.094517	21	21
5	SHYLOCK	5	18	6	20	0.090521	20	19.33333333
6	ANTONIO	6	20	5	18	0.063797	19	19
7	GRATIANO	5	18	5	18	0.061303	18	18
8	LAUNCELOT	4	17	3	16	0.056909	16	16.33333333
9	DUKE	2	13	4	17	0.05781	17	15.66666667
10	JESSICA	3	15	2	13	0.043711	15	14.33333333
11	NERISSA	2	13	2	13	0.042869	14	13.33333333
12	SALERIO	3	15	1	6	0.042207	13	11.33333333
13	SALARINO	1	6	2	13	0.024489	12	10.33333333
14	GOBBO	1	6	1	6	0.016195	5	5.666666667
15	MOROCCO	0	1	1	6	0.016472	7	4.666666667
16	ARRAGON	0	1	1	6	0.016472	7	4.666666667
17	Servant	1	6	0	1	0.016472	7	4.666666667
18	ALL	1	6	0	1	0.016472	7	4.666666667
19	BALTHASAR	1	6	0	1	0.016472	7	4.666666667
20	Clerk	0	1	1	6	0.016352	6	4.333333333
21	SALANIO	0	1	1	6	0.016142	3	3.333333333
22	TUBAL	1	6	0	1	0.016142	3	3.333333333
23	STEPHANO	0	1	1	6	0.01545	2	3
24	LEONARDO	1	6	0	1	0.015305	1	2.666666667

Table 23:*7.2.7 Much Ado about Nothing Mentioning*

1	Id	In-Degree	RANK	Out-Degree	RANK	PageRank	RANK	AVERAGE RANK
2	DONPEDRO	6	17	7	20	0.090066	18	18.33333333
3	BENEDICK	6	17	6	18	0.101595	20	18.33333333
4	LEONATO	6	17	4	14	0.090225	19	16.66666667
5	CLAUDIO	6	17	4	14	0.079262	17	16
6	BEATRICE	3	14	6	18	0.075958	14	15.33333333
7	HERO	3	14	5	17	0.077203	15	15.33333333
8	DOGBERRY	3	14	4	14	0.078711	16	14.66666667
9	VERGES	1	2	3	12	0.050383	13	9
10	FRIARFRANCIS	2	10	1	6	0.040002	10	8.666666667
11	DONJOHN	1	2	3	12	0.043219	11	8.333333333
12	BORACHIO	1	2	2	11	0.04779	12	8.333333333
13	CONRADE	2	10	1	6	0.034452	9	8.333333333
14	ANTONIO	2	10	1	6	0.031834	7	7.666666667
15	MARGARET	2	10	0	1	0.029218	6	5.666666667
16	URSULA	1	2	1	6	0.031947	8	5.333333333
17	Lord	0	1	1	6	0.01873	3	3.333333333
18	Sexton	1	2	0	1	0.021783	5	2.666666667
19	Watchman	1	2	0	1	0.020898	4	2.333333333
20	BALTHASAR	1	2	0	1	0.018436	2	1.666666667
21	Boy	1	2	0	1	0.018289	1	1.333333333

Table 24:*7.2.8 Othello Mentioning*

1	Id	In-Degree	RANK	Out-Degree	RANK	PageRank	RANK	AVERAGE RANK
2	IAGO	10	14	10	15	0.1392313	14	14.33333333
3	OTHELLO	10	14	7	14	0.1462442	15	14.33333333
4	CASSIO	5	12	6	13	0.0955371	13	12.66666667
5	DESDEMONA	5	12	4	11	0.0779858	12	11.66666667
6	LODOVICO	4	10	3	8	0.0649618	11	9.666666667
7	GRATIANO	1	4	4	11	0.0534472	9	8
8	BRABANTIO	4	10	1	2	0.0581178	10	7.333333333
9	EMILIA	3	8	3	8	0.0417814	6	7.333333333
10	RODERIGO	3	8	2	4	0.0431934	7	6.333333333
11	DUKEOFVENICE	0	1	3	8	0.0444116	8	5.666666667
12	BIANCA	2	6	2	4	0.0308012	4	4.666666667
13	MONTANO	2	6	2	4	0.0305872	3	4.333333333
14	FirstSenator	0	1	2	4	0.0332173	5	3.333333333
15	Clown	1	4	0	1	0.0208251	2	2.333333333
16	Herald	0	1	1	2	0.0206111	1	1.333333333

Table 25:*7.2.9 Romeo Juliet Mentioning*

1	Id	In-Degree	RANK	Out-Degree	RANK	PageRank	RANK	AVERAGE RANK
2	ROMEO	8	26	9	25	0.10358	26	25.66666667
3	BENVOLIO	7	25	5	21	0.086059	24	23.33333333
4	CAPULET	4	18	9	25	0.090422	25	22.66666667
5	Nurse	6	22	6	24	0.063456	22	22.66666667
6	PRINCE	6	22	5	21	0.075098	23	22
7	JULIET	6	22	4	18	0.053246	20	20
8	FRIARLAURENCE	4	18	5	21	0.056716	21	20
9	LADYCAPULET	5	21	3	15	0.052416	19	18.33333333
10	MERCUTIO	4	18	4	18	0.03732	18	18
11	TYBALT	3	17	4	18	0.037311	17	17.33333333
12	MONTAGUE	2	13	3	15	0.030746	14	14
13	Servant	2	13	2	12	0.036806	16	13.66666667
14	BALTHASAR	2	13	2	12	0.029579	13	12.66666667
15	FirstWatchman	1	4	3	15	0.034183	15	11.33333333
16	SAMPSON	0	1	2	12	0.028834	12	8.333333333
17	LADYMONTAGUE	2	13	0	1	0.022391	11	8.333333333
18	PARIS	1	4	1	7	0.020786	10	7
19	SecondWatchman	1	4	1	7	0.015234	8	6.333333333
20	FirstCitizen	1	4	1	7	0.013684	7	6
21	GREGORY	1	4	0	1	0.017816	9	4.666666667
22	Chorus	0	1	1	7	0.013563	3	3.666666667
23	NURSE	0	1	1	7	0.013563	3	3.666666667
24	FirstMusician	1	4	0	1	0.01359	6	3.666666667
25	Apothecary	1	4	0	1	0.013563	3	2.666666667
26	PETER	1	4	0	1	0.013258	2	2.333333333
27	SecondCapulet	1	4	0	1	0.013244	1	2

Table 26:*7.2.10 Taming of the Shrew Mentioning*

1	Id	In-Degree	RANK	Out-Degree	RANK	PageRank	RANK	AVERAGE RANK
2	PETRUCHIO	9	30	11	30	0.06262	29	29.66666667
3	TRANIO	8	29	9	29	0.043762	25	27.66666667
4	BAPTISTA	7	28	6	26	0.042907	24	26
5	HORTENSIO	5	23	7	27	0.045235	26	25.33333333
6	LUCENTIO	4	19	8	28	0.046692	27	24.66666667
7	SLY	4	19	5	22	0.068258	30	23.66666667
8	Lord	4	19	4	21	0.052335	28	22.66666667
9	VINCENTIO	5	23	5	22	0.029691	21	22
10	GREMIO	5	23	5	22	0.029657	20	21.66666667
11	GRUMIO	5	23	3	19	0.034189	22	21.33333333
12	KATHARINA	4	19	5	22	0.026628	15	18.66666667
13	BIONDELLO	5	23	2	15	0.025331	14	17.33333333
14	BIANCA	3	17	3	19	0.025176	13	16.33333333
15	Pedant	3	17	2	15	0.020791	10	14
16	NATHANIEL	1	2	2	15	0.041998	23	13.33333333
17	FirstServant	2	16	1	6	0.027241	16	12.66666667
18	FirstHuntsman	1	2	2	15	0.022861	11	9.333333333
19	JOSEPH	1	2	1	6	0.028229	19	9
20	PHILIP	1	2	1	6	0.028033	17	8.333333333
21	NICHOLAS	1	2	1	6	0.028033	17	8.333333333
22	SecondHuntsman	1	2	1	6	0.022861	11	6.333333333
23	ThirdServant	1	2	1	6	0.013871	7	5
24	APlayer	1	2	1	6	0.0131	6	4.666666667
25	CURTIS	1	2	1	6	0.009999	4	4
26	PETER	1	2	0	1	0.016047	9	4
27	Tailor	0	1	1	6	0.009999	4	3.666666667
28	SecondServant	1	2	0	1	0.013871	7	3.333333333
29	Servant	1	2	0	1	0.009534	3	2
30	Haberdasher	1	2	0	1	0.008273	1	1.333333333
31	Widow	1	2	0	1	0.008273	1	1.333333333

Table 27:*7.2.11 Tempest Mentioning*

1	Id	In-Degree	RANK	Out-Degree	RANK	PageRank	RANK	AVERAGE RANK
2	PROSPERO	5	16	8	16	0.134357	16	16
3	ALONSO	4	15	5	15	0.075185	15	15
4	SEBASTIAN	3	9	3	12	0.061269	13	11.33333333
5	STEPHANO	3	9	3	12	0.059323	11	10.66666667
6	GONZALO	3	9	2	7	0.060187	12	9.333333333
7	CERES	2	6	2	7	0.061325	14	9
8	Boatswain	3	9	2	7	0.056396	10	8.666666667
9	ANTONIO	1	1	3	12	0.050202	8	7
10	TRINCULO	2	6	2	7	0.047018	7	6.666666667
11	FERDINAND	3	9	1	1	0.051775	9	6.333333333
12	CALIBAN	3	9	1	1	0.045809	6	5.333333333
13	IRIS	1	1	2	7	0.036933	5	4.333333333
14	MIRANDA	2	6	1	1	0.034565	4	3.666666667
15	JUNO	1	1	1	1	0.025401	3	1.666666667
16	Master	1	1	1	1	0.024105	2	1.333333333
17	ARIEL	1	1	1	1	0.019547	1	1

Table 28:*7.2.12 Twelfth Night Mentioning*

1	Id	In-Degree	RANK	Out-Degree	RANK	PageRank	RANK	AVERAGE RANK
2	OLIVIA	5	14	5	14	0.107835	14	14
3	VIOLA	4	11	3	9	0.093099	12	10.66666667
4	SIRTOBYBELCH	2	7	4	12	0.099039	13	10.66666667
5	MALVOLIO	4	11	3	9	0.092952	11	10.33333333
6	SEBASTIAN	3	10	4	12	0.071922	8	10
7	ANTONIO	2	7	3	9	0.076272	10	8.666666667
8	FABIAN	4	11	1	4	0.074511	9	8
9	DUKEORSINO	2	7	2	6	0.061271	7	6.666666667
10	Clown	1	3	2	6	0.05719	6	5
11	MARIA	0	1	2	6	0.042562	5	4
12	CURIO	1	3	0	1	0.027008	4	2.666666667
13	SIRANDREW	1	3	0	1	0.02658	3	2.333333333
14	SecondOfficer	0	1	1	4	0.025923	2	2.333333333
15	Captain	1	3	0	1	0.025662	1	1.666666667

7.3 Time Series Analysis:

Following are more examples of Time series analysis

7.3.1 As you like it

There are total 5 ACTs in the play with multiple number of SCENEs in each ACT.

7.3.1.1 Interaction

The list of important characters considered for the time series analysis for the play *As you like it*– interaction edge list are

1) ORLANDO

2) ROSALIND

3) CELIA

4) OLIVER

5) TOUCHSTONE

6) JAQUES

The Cytoscape networks are created for each scene of the interaction edge list of *As you like it* after the actual edge list is mined to retrieve the edge list confined only to the important characters listed above.

Note: Network analysis is done by mapping node size to Degree Centrality and Node Color to Betweenness Centrality.

Figure 11:*7.1 As you like it Interaction - ACT I*

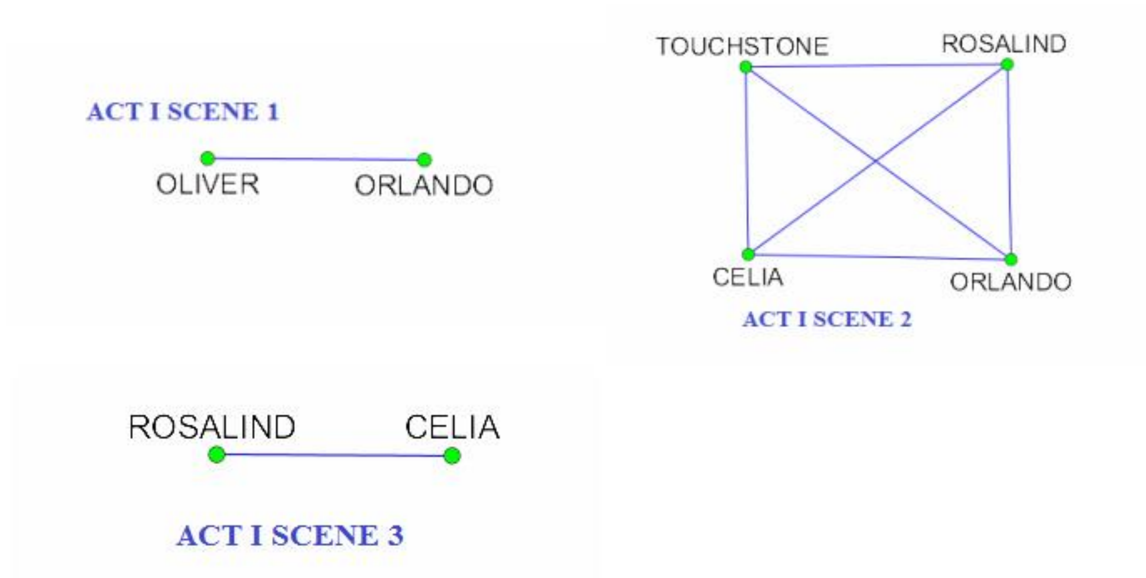


Figure 12:*7.2 As you like it Interaction - ACT II*

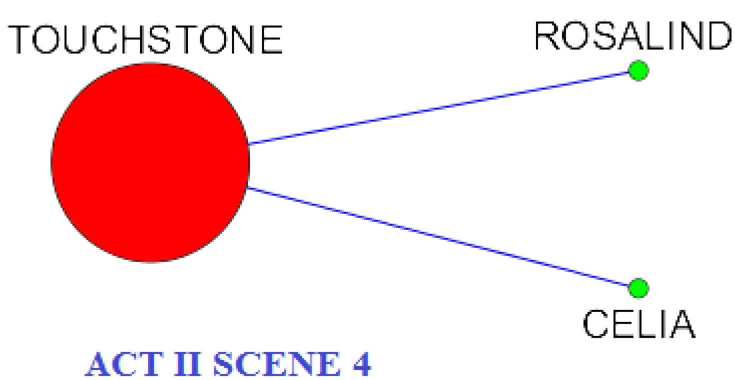


Figure 13: *7.3 As you like it Interaction - ACT III*

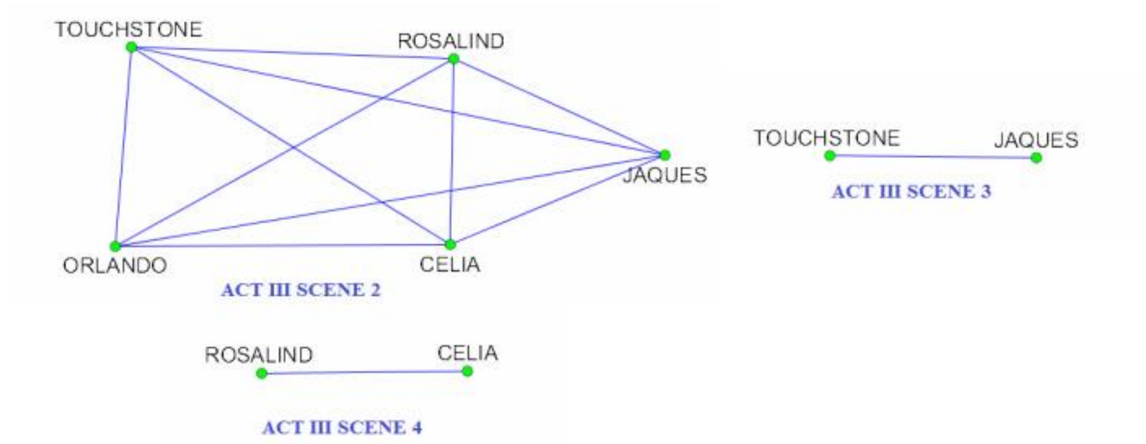


Figure 14: *7.4 As you like it Interaction - ACT IV*

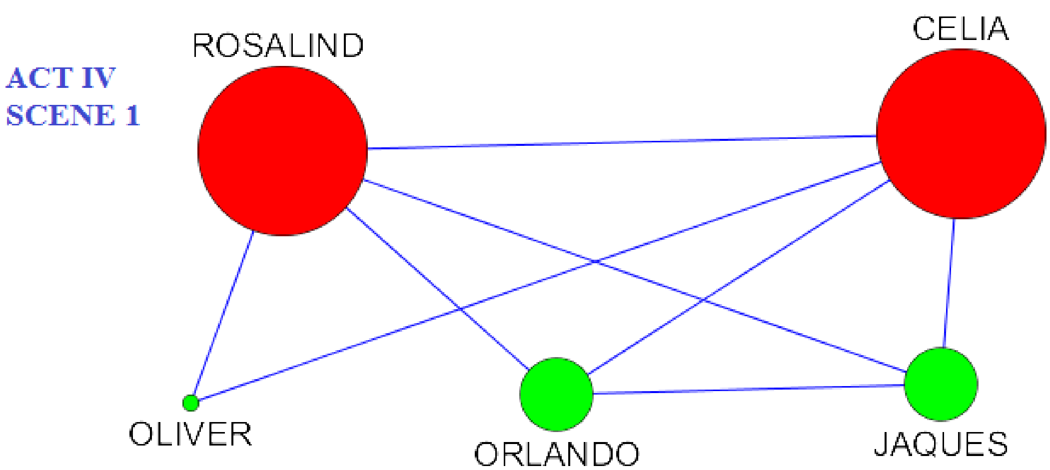
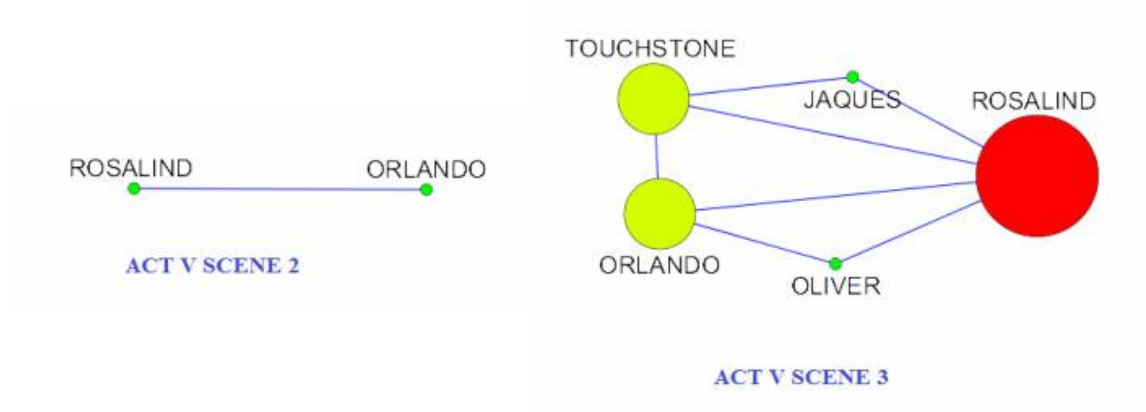


Figure 15:*7.5 As you like it Interaction - ACT V*



7.3.1.2 Mentioning

The list of important characters considered for the time series analysis for the play as you like – Mentioning edge list are

1) ORLANDO
2) ROSALIND
3) TOUCHSTONE
4) OLIVER
5) DUKESENIOR
6) JAQUES

Figure 16: *7.6 As you like it Mentioning - ACT I*

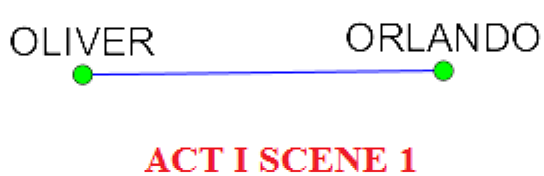


Figure 17:*7.7 As you like it Mentioning - ACT III*

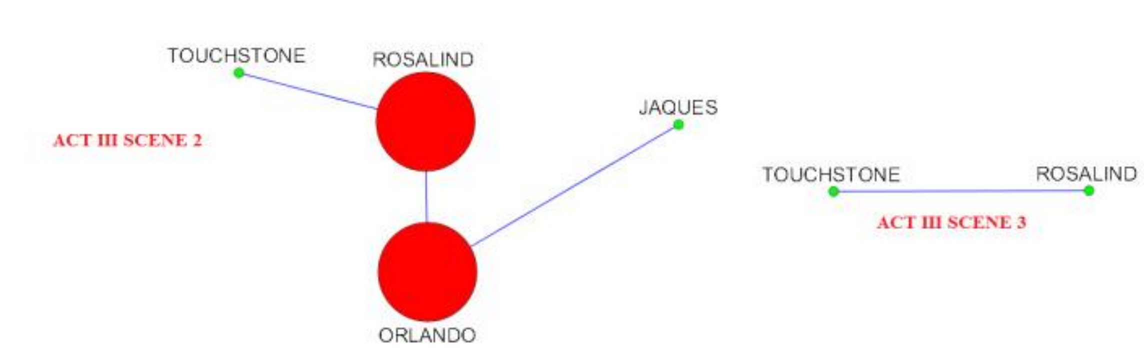


Figure 18: *7.8 As you like it Mentioning - ACT IV*

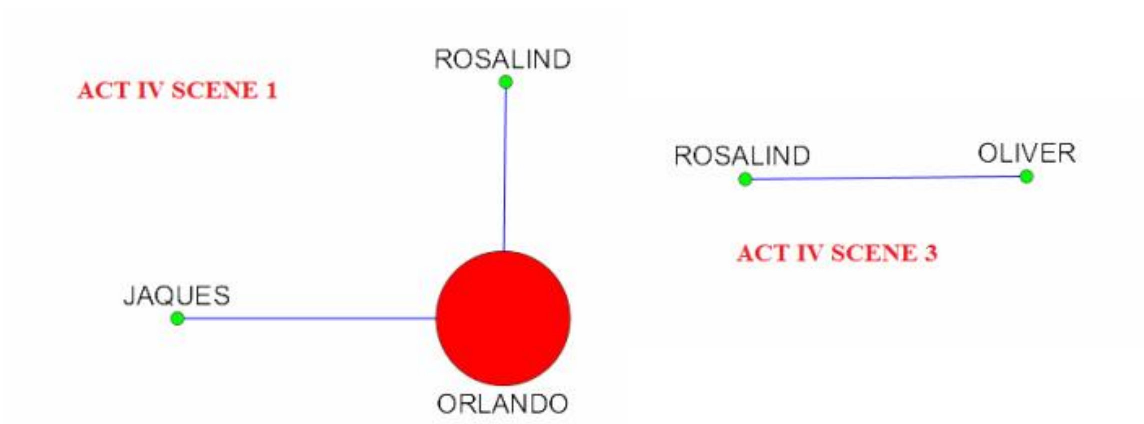
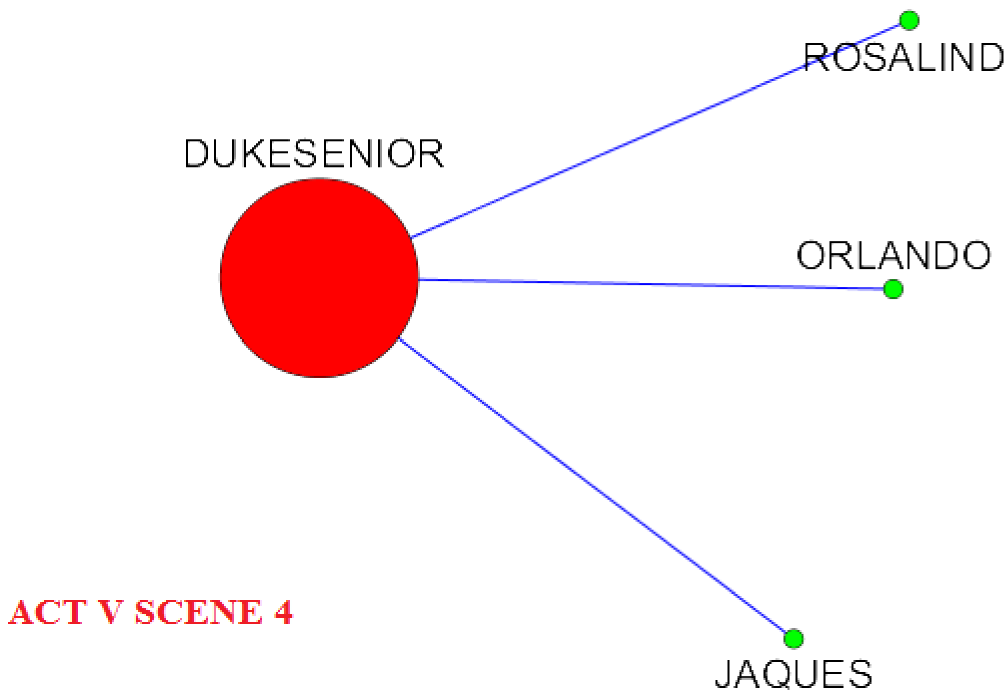


Figure 19: *7.9 As you like it Mentioning - ACT V*



7.3.2 Hamlet

7.3.2.1 Mentioning

The list of important characters considered for the time series analysis for the play

Hamlet – Mentioning edge list are

1) KINGCLAUDIUS
2) HAMLET
3) HORATIO
4) LAERTES
5) BERNARDO
6) OPHELIA
7) QUEENGERTRUDE

Figure 20: *7.10 Hamlet Mentioning - ACT I*

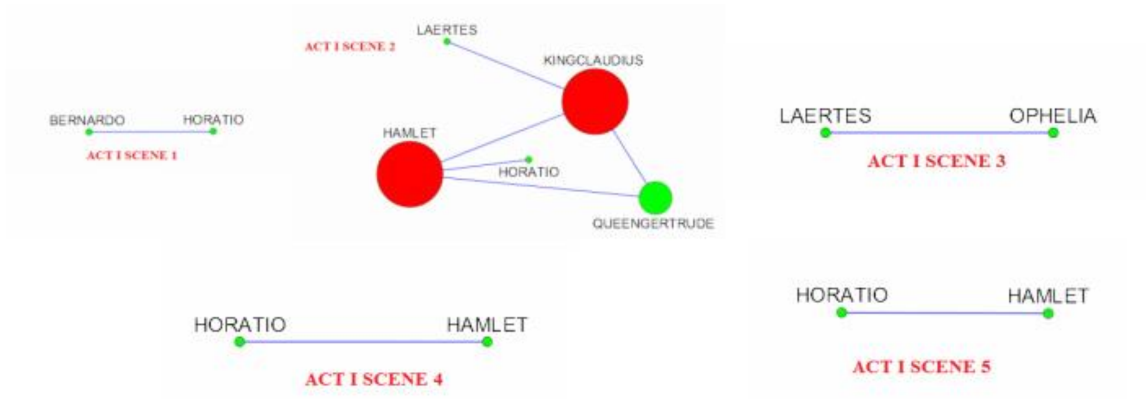


Figure 21: *7.11 Hamlet Mentioning - ACT II*

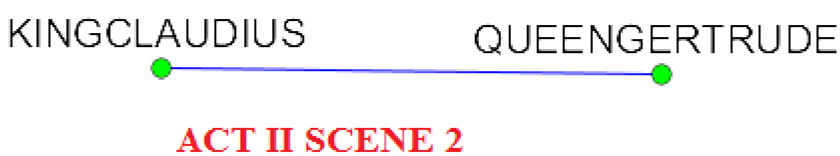


Figure 22: *7.12 Hamlet Mentioning - ACT III*

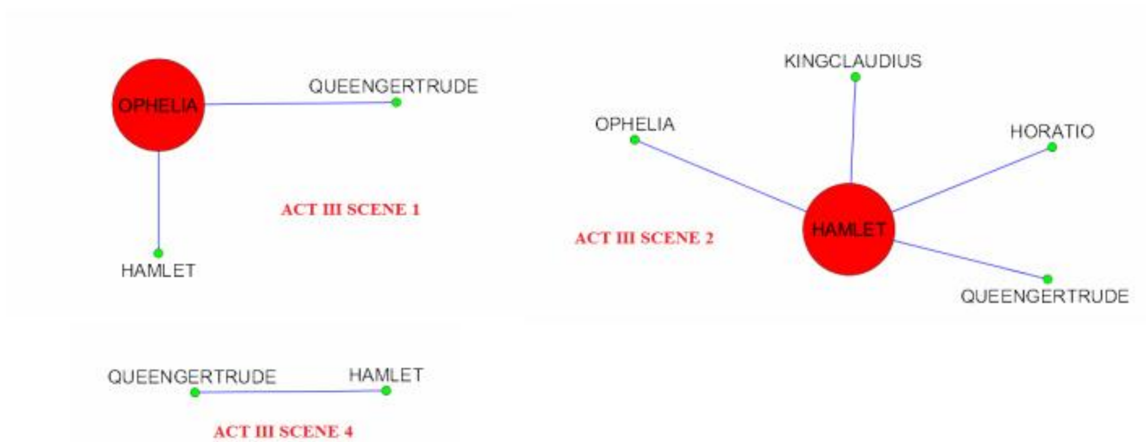


Figure 23:*7.13 Hamlet Mentioning - ACT IV*

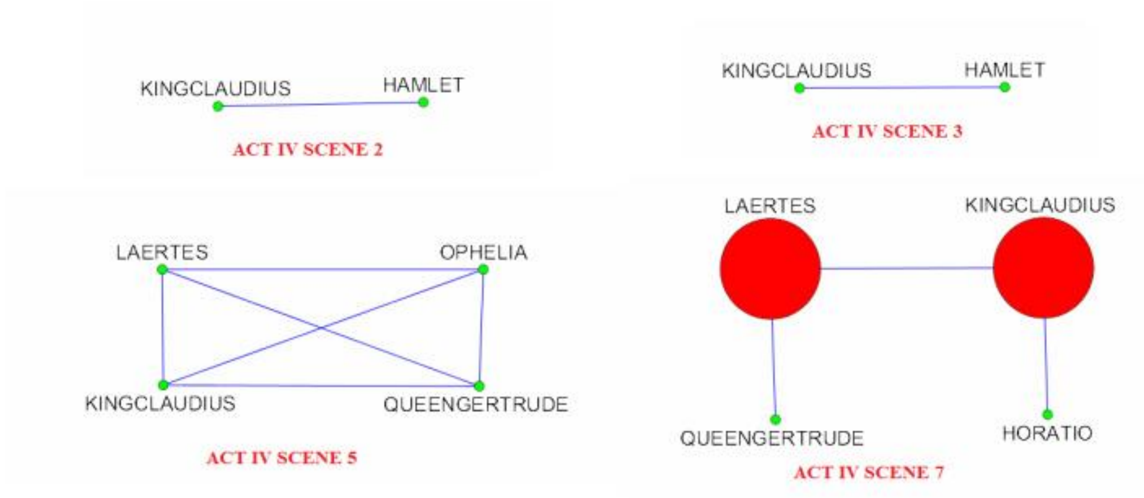


Figure 24:*7.14 Hamlet Mentioning - ACT V*

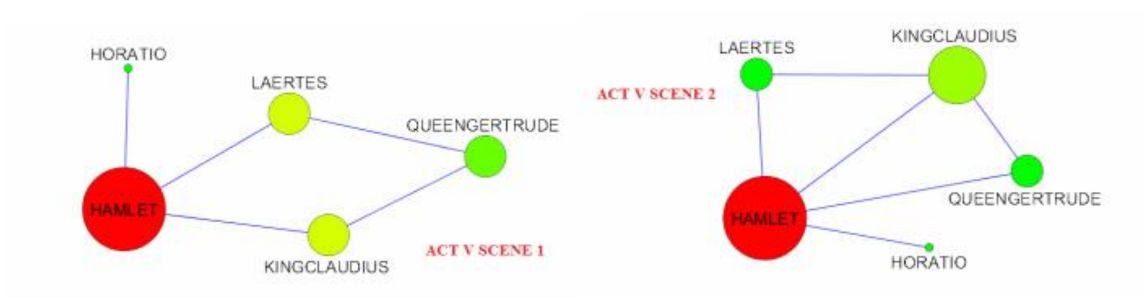


Table 29: *7.15 Gephi Analysis*

Julius Caesar	**Interaction**				**Mentioning**		
Characters	*Degree*	*Betweenness*	*Closeness*	*Eigenvector*	*In degree*	*Out degree*	*Page Rank*
Brutus	*High*	*High*	*Low*	*High*	*High*	*High*	*High*
Cassius	*High*	*Low*	*Low*	*High*	*High*	*High*	*High*
Antony	*High*	*High*	*Low*	*High*	*High*	*Low*	*High*
Servant	*Low*	*Low*	*Low*	*High*	*Low*	*Low*	*Low*
Casca	*Low*	*Low*	*Low*	*High*	*Low*	*Low*	*Low*

Caesar	*Low*	*Low*	*High*	*High*	*High*	*High*	*High*
King Lear	**Interaction**				**Mentioning**		
Characters	*Degree*	*Betweenness*	*Closeness*	*Eigenvector*	*In degree*	*Out degree*	*Page Rank*
Goneril	*High*	*High*	*Low*	*High*	*High*	*Low*	*High*
Gloucester	*High*	*High*	*High*	*High*	*High*	*Low*	*High*
Regan	*High*	*High*	*Low*	*High*	*Low*	*Low*	*High*
Kent	*High*	*Low*	*High*	*High*	*High*	*Low*	*High*
Edgar	*High*	*Low*	*High*	*High*	*Low*	*Low*	*Low*
King Lear	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Edmund	*High*	*Low*	*High*	*High*	*High*	*Low*	*High*
Macbeth	**Interaction**				**Mentioning**		
Characters	*Degree*	*Betweenness*	*Closeness*	*Eigenvector*	*In degree*	*Out degree*	*Page Rank*
Macbeth	*High*	*High*	*Low*	*High*	*High*	*High*	*High*
Lennox	*High*	*High*	*High*	*High*	*Low*	*Low*	*Low*
Banquo	*High*	*Low*	*High*	*High*	*Low*	*Low*	*Low*
Ross	*High*	*High*	*High*	*High*	*Low*	*Low*	*Low*
Lady Macbeth	*High*	*Low*	*High*	*High*	*Low*	*Low*	*Low*
Malcom	*High*	*Low*	*High*	*High*	*Low*	*Low*	*Low*
Macduff	*Low*	*Low*	*High*	*High*	*Low*	*Low*	*Low*
First Witch	*Low*	*Low*	*High*	*High*	*Low*	*Low*	*Low*
Merchant of Venice	**Interaction**				**Mentioning**		

Characters	Degre e	Betwee nness	Closenes s	Eigenvector	In degree	Out degree	Page Rank
Portia	*High*	*High*	*Low*	*High*	*High*	*High*	*High*
Gratiano	*High*	*High*	*High*	*High*	*High*	*High*	*High*
Nerissa	*High*	*Low*	*High*	*High*	*Low*	*Low*	*Low*
Bassanio	*High*	*High*	*High*	*High*	*High*	*High*	*High*
Lorenzo	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Shylock	*Low*	*High*	*High*	*High*	*High*	*High*	*High*
Antonio	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Much Ado About Nothing	**Interaction**				**Mentioning**		
Characters	*Degre e*	*Betwee nness*	*Closenes s*	*Eigenvector*	*In degree*	*Out degree*	*Page Rank*
Leonato	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Don Pedro	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Claudio	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Borachio	*High*	*High*	*High*	*High*	*Low*	*Low*	*Low*
Benedick	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Hero	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Beatrice	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Othello	**Interaction**				**Mentioning**		
Characters	*Degre e*	*Betwee nness*	*Closenes s*	*Eigenvector*	*In degree*	*Out degree*	*Page Rank*
Iago	*High*	*High*	*High*	*High*	*High*	*High*	*High*
Othello	*High*	*High*	*High*	*High*	*High*	*High*	*High*
Desdemo	*High*	*High*	*High*	*High*	*High*	*Low*	*High*

na							
Roderigo	*High*	*Low*	*High*	*High*	*Low*	*Low*	*Low*
Cassio	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Romeo Juliet	**Interaction**				**Mentioning**		
Characters	*Degre e*	*Betwee nness*	*Closenes s*	*Eigenvector*	*In degree*	*Out degree*	*Page Rank*
Romeo	*High*	*High*	*High*	*High*	*High*	*High*	*High*
Capulet	*High*	*High*	*Low*	*High*	*High*	*High*	*High*
Lady Capulet	*High*	*High*	*Low*	*High*	*High*	*Low*	*High*
Benvolio	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Prince	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Nurse	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Taming of the Shrew	**Interaction**				**Mentioning**		
Characters	*Degre e*	*Betwee nness*	*Closenes s*	*Eigenvector*	*In degree*	*Out degree*	*Page Rank*
Petruchio	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Katharin a	*High*	*High*	*High*	*High*	*Low*	*Low*	*Low*
First Servant	*High*	*High*	*High*	*High*	*Low*	*Low*	*Low*
Grumio	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Tranio	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Baptista	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Hortenisi o	*High*	*Low*	*High*	*High*	*High*	*High*	*High*

Tempest	Interaction				Mentioning		
Characters	*Degree*	*Betweenness*	*Closeness*	*Eigenvector*	*In degree*	*Out degree*	*Page Rank*
Prospero	*High*	*High*	*High*	*High*	*High*	*High*	*High*
Ariel	*High*	*High*	*High*	*High*	*Low*	*Low*	*Low*
Sebastian	*High*	*Low*	*High*	*High*	*High*	*Low*	*High*
Antonio	*High*	*Low*	*High*	*High*	*Low*	*Low*	*High*
Alonso	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Stephano	*High*	*Low*	*High*	*High*	*High*	*Low*	*High*
Gonzalo	*High*	*Low*	*High*	*High*	*High*	*Low*	*High*
Twelfth Night	**Interaction**				**Mentioning**		
Characters	*Degree*	*Betweenness*	*Closeness*	*Eigenvector*	*In degree*	*Out degree*	*Page Rank*
Viola	*High*	*High*	*Low*	*High*	*High*	*High*	*High*
Sir Toby Belch	*High*	*Low*	*High*	*High*	*Low*	*High*	*High*
Sir Andrew	*High*	*Low*	*High*	*High*	*Low*	*Low*	*Low*
Clown	*High*	*Low*	*High*	*High*	*Low*	*Low*	*High*
Olivia	*High*	*Low*	*High*	*High*	*High*	*High*	*High*
Malvolio	*High*	*Low*	*High*	*High*	*High*	*High*	*High*

Chapter 8

References

[1] *Inferring and Visualizing Social Networks on Internet Relay Chat*

By Paul Mutton {pjm2@kent.ac.uk}, Computing Laboratory, University of Kent,

Canterbury, Kent.CT2 7NF. United Kingdom.

Information Visualization, 2004. IV 2004. Proceedings. Eighth International

Conference on 14-16 July 2004

[2] *THE SMALL WORLD OF SHAKESPEARE'S PLAYS*

James Stiller, University of Gloucestershire,

Daniel Nettle, The Open University and

Robin I. M. Dunbar, University of Liverpool.

Human Nature, December 2003, Volume 14, Issue 4, pp 397-408

[3] *Social Networks Generate Interest in Computer Science*

Casey Alt, Owen Astrachan, Jeffrey Forbes,Richard Lucic, Susan Rodger

Department of Computer Science Duke University, Durham,NC 27708

SIGCSE '06 Proceedings of the 37th SIGCSE technical symposium on Computer

science education

[4] *The performance of modularity maximization in practical contexts*

B. H. Good, Y.-A.deMontjoye and A. Clauset

Physical Review. E, 81, 046106 (2010). Published 15 April 2010

[5] *http://www.cytoscape.org/*

[6] *http://gephi.github.io/*

[7] *http://www.activatenetworks.net/who-is-central-to-a-social-network-it-depends-on-your-centrality-measure/*

[8] *On Variants of Shortest-Path Betweenness Centrality and their Generic Computation*

Ulrik Brandes, Department of Computer & Information Science, University of Konstanz

Social Networks, Volume 30, Issue 2, May 2008, Pages 136–145

[9] *Network Centrality*

Based on materials by Lada Adamic, UMichigan

[10] *Centrality in Social Networks Conceptual Clarification*

Linton C. Freeman, Lehigh University,

Social Networks Volume 1, Issue 3, 1978–1979, Pages 215–239

[11] *Scale-Free Characteristics of Random Networks: The Topology of the World-Wide Web*

Barabasi, A.-L., R. Albert, and H. Jong

[12] *Why Fiction May Be Twice as True as Fact: Fiction as Cognitive and Emotional Simulation*

Keith Oatley, University of Toronto

Physica A: Statistical Mechanics and its Applications

Volume 281, Issues 1–4, 15 June 2000, Pages 69–77

[13] *Learning Relations among Movie Characters: A Social Network Perspective*

Lei Ding and Alper Yilmaz,

Programmatic Computer vision lab, The Ohio State University

Computer Vision – ECCV 2010

Lecture Notes in Computer Science Volume 6314, 2010, pp 410-423

[14] *Gephi: An Open Source Software for Exploring and Manipulating Networks*

Mathieu Bastian and Sebastien Heymann, Mathieu Jacomy

Third International AAAI Conference on Weblogs and Social Media

May 17, 2009 – May 20, 2009

[15] [BOOK] *Statistical Analysis of Network Data*

[16] *Network Theory, Plot Analysis*

Franco Moretti

[17] *Extracting Sentiment Networks from Shakespeare's Plays*

Eric T. Nalisnick and Henry S. Baird

Dept. of Computer Science and Engineering
Lehigh University
Document Analysis and Recognition (ICDAR), 2013 12th International Conference on

25-28 Aug. 2013

[18] *http://shakespeare.mit.edu/*

[19]*http://www.william-shakespeare.info/william-shakespeare-biography.htm*