Revista de Sistemas de Informação da FSMA
n. 27 (2021) pp. 54-66 http://www.fsma.edu.br/si/sistemas.html

Comparing the Centralities of Core Characters in Three Sitcoms

Ana L.C. Bazzan, *Instituto de Informática – UFRGS, Porto Alegre, Brazil*

***Abstract*—The area of network science provides tools to analyze the structure of the social network of characters underlying the plot of narratives such as sitcoms. It can also help to detect roles and characters that are prominent. In this paper, three sitcoms – *Seinfeld*, *Friends*, and *The Big Bang Theory*– are analyzed, whose plots are centered on a couple of friends who face typical daily situations. For such analysis, characteristic measures of their social networks are used, as well as measures that compute the importance of single characters, when interacting with others in the network. The contributions of the present paper are manifold. For one side, it analyzes social networks that resemble social relations among young people. A second contribution is towards extending previous studies that compare those sitcoms. Lastly, this paper also investigates whether there is a correlation between the respective ratings and a measure of the distribution of degrees.**

***Index Terms*—Social networks, Centrality measures, Sitcom.**

I. Introduction

What makes some sitcoms more popular than others? Apart from obvious candidate answers such as plot, cast, and cinematography, the area of network science also tries to contribute to answering that question by providing tools to analyze the structure of the network of characters underlying the plot (the social network of the show's plot), as well as to detect roles that are prominent.

Thus, analyzing individual shows or comparing similar ones seem to make sense, as it was done in [19] (a survey paper on this subject), as well as in [9], [10], [11], [16], [17], [18].

In this paper, three sitcoms – *Seinfeld*, *Friends*, and *The Big Bang Theory* – are analyzed, whose plots are centered on a couple of friends who face typical daily situations for those who live in a big city, even if one show is known to be about nothing. For such analysis, typical measures of their social networks are used, as well as measures that compute the importance of single characters, when interacting with others in the network. Examples of such social networks appear in figures 1 to 7.

While a previous work [10] has proposed a similar comparison, that work covers only a subset of episodes. The contributions of the present paper are to extend that comparison to cover the majority of the seasons (if not all) of the three aforementioned sitcoms. Another contribution is to collect data about the ratings of the shows and analyze whether they correlate with a measure of the distribution of degrees.

This paper is organized as follows: the next section discusses other works that deal with tools of network science for quantifying some aspects of storytelling. Sections III to V give brief descriptions of the three shows. Section VI discusses the structure of the social networks of the shows, from a global network perspective. A perspective centered on the centrality of the respective characters is the focus of Section VII. Section VIII presents some concluding remarks.

II. Storytelling and social networks

In this section, the benefits of using tools of social networks to analyze the structure of well-known sitcoms and other shows is briefly discussed. The reader is referred to [19] for more details, as well as pointers to related works.

Beveridge and Shan [11] used network theory to investigate who is/are the most central characters in *Game of Thrones*. This popular show was also the target of: [20], [29] and [18]. The latter computed the importance of characters and used them as features or input to a machine learning algorithm in order to predict how likely to die some characters are.

The three sitcoms featured in the present work have also received attention in past works: [30] (featuring *Seinfeld*), [13] (*The Big Bang Theory*), while *Friends* was the focus of several works, as follows. Nan et al. [22] used a deep learning model for face recognition in *Friends*'s videos in order to distinguish the six main characters and establish the social network between them. Albright [3] calculated the frequency of characters' shared plotlines, throughout the entire show, drawing conclusions on who are the most independent characters. Analyzing the importance of the characters is also the goal of [28]. Seth [27] used the transcripts of *Friends* available in the Internet, aiming at shedding light on the question about who stood out among the character of the show. The following parameters for each character were accounted for: number of lines and words spoken, number of screen appearances, appearances in some locations, and mentions in the episode title. A similar goal underlies the work in [8], [9], [10], where the importance of each character is investigated using various centrality measures. In particular, in [10] four sitcoms are quantitatively compared, showing that despite an intuition that they are very similar, such intuition cannot be backed by the centrality measure values.

Still focusing on *Friends*, [17], [16] compared different extraction methods using both manually extracted and automated datasets, providing evidence that automated methods of data extraction as, e.g., machine learning, are reliable for most (though not all) analyses.

Regarding *The Big Bang Theory*, a recent paper [13] has focused on the dialogues of the sitcom *The Big Bang Theory*, using a concentration measure to analyze dominance in the dialogues. The authors show a declining trend in the concentration of dialogues over the seasons. Their main finding is that there is a high correlation between the decline in that concentration and a decline in popularity of the show. This has inspired the investigation that is discussed ahead (Section VI), which concludes that, when one measures the concentration by means of entropy of degree distribution, one no longer observes such correlation.

Other works, dealing with extraction of interactions (from video, transcripts, etc.) are for instance [21], [12].

These works all investigate specific aspects of individual shows; however only a few aim at comparing them. Among them, we can highlight the following: (i) [31], where the authors have analyzed the character networks of *Stargate* and *Star Trek* and found that their structures are similar; (ii) [10], as aforementioned. Hence, there is a gap in the literature regarding comparison of shows that, intuition says, look similar.

III. The Structure of the Social Network of *Seinfeld*

Created by Larry David and Jerry Seinfeld for NBC, *Seinfeld* features Jerry Seinfeld (himself), his school friend George Costanza (Jason Alexander), his former girlfriend Elaine Benes (Julia Louis-Dreyfus), and his neighbor across the hall Cosmo Kramer

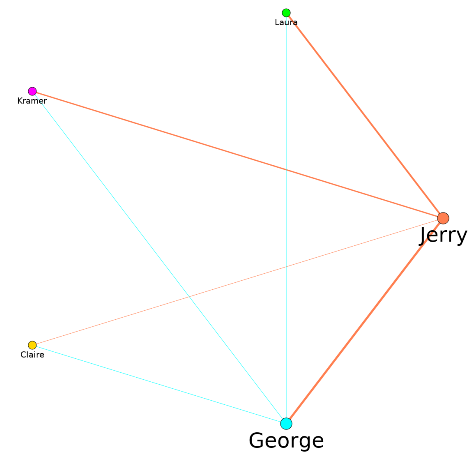

Fig. 1: *Seinfeld*: Network of Characters (pilot).

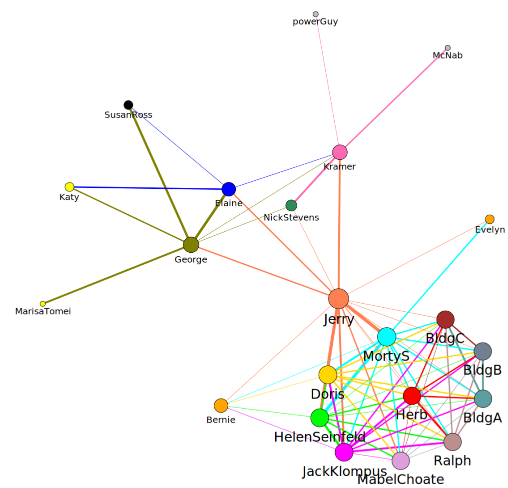

Fig. 2: *Seinfeld*: Network of Characters (episode with big clique).

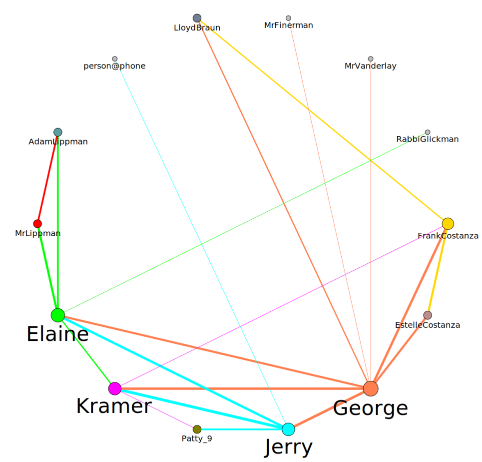

Fig. 3: *Seinfeld*: Network of Characters of episode "The Serenity Now".

(Michael Richards). The show is set predominantly in an apartment building in Manhattan. The "show about nothing"—as often described—began without much fuss in 1989, with a pilot and only 4 other episodes in Season 1, while Season 2 had 12 episodes. After, Season 3 had 23 episodes[1] and, after that, it became one of the biggest comedy hits in the U.S.

The stories are about the minutiae of daily life, as experienced by four thirty-something single New Yorkers who had no family or other strong responsibilities and hence, allow room for obsessions about small things such as getting a table in a Chinese restaurant, queuing, renting an apartment, finding the car in a parking garage, buying a new suit, getting together with friends, etc.

Since its beginning, *Seinfeld* broke several sitcom structures and formulas such a central romantic relationship: Larry David is credited with refusing to focus on a romantic relationship formula between Jerry and Elaine. Rather, episodes would follow a proper structure: the story thread is presented at the beginning, normally involving the characters starting in their own situations; this is then followed by rapid scene-shifts between plot lines bringing the stories together. Thus the characters' stories intertwine in each episode, and, despite the separate plot lines, the narratives maintain the ties among the four characters. This means a change over the usual A and B plotlines. Also, *Seinfeld* departed from family and group sitcom formulas of its time, in that the four main characters were not related by family or work connections, but remain close friends throughout the episodes.

Another characteristic of *Seinfeld* is that a higher than usual number of secondary characters recur and, moreover, play an important role, at least at single episode level. In fact, as discussed ahead, the number of characters throughout the episodes in *Seinfeld* is much higher than the other sitcoms. These characteristics have obvious consequences for the social network structure of the show; characters in *Seinfeld* have the lowest average degree (number of interactions they are involved in), a clear consequences of the aforementioned separated but tied plotlines. It might be also that the scenes in *Seinfeld* are longer and thus there are less connections.

In order to see how the structure of the show has changed, see Fig. 1, Fig. 2, and Fig. 3. The former shows the network of the very first episode of Seinfeld (where one clearly sees the prominent role of Jerry, followed by George; Kramer interacts with both only; Elaine did not take part). The latter depicts the social network of an episode with the maximum clique (Episode 14 in Season 7, where a bunch of characters living in the same condo as Jerry's parents take place in a meeting of the condo's board of the directors). The network in Fig. 3 is closer to a typical one.

[1]The number of episodes per season varies from source to source, given that some episodes were aired together or not, depending on the country or media (DVD, streaming, etc.).

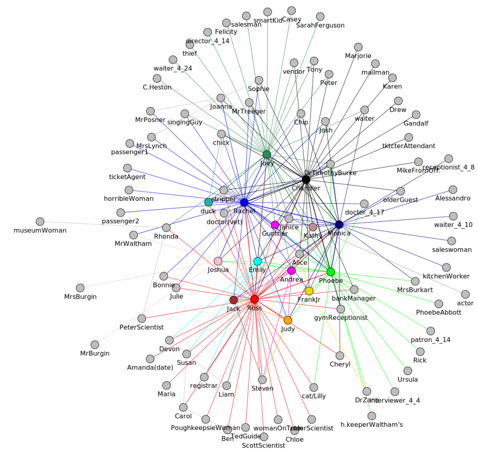

Fig. 4: *Friends*: Network of Characters (Season Four).

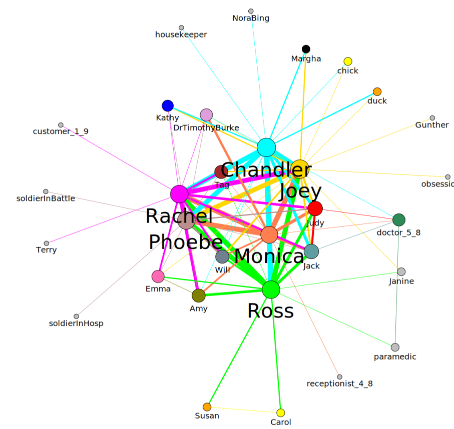

Fig. 5: *Friends*: Network of Characters (Thanksgiving episodes).

IV. The Structure of the Social Network of *Friends*

Friends is an American television sitcom created by David Crane and Marta Kauffman, which was aired on NBC from 1994 to 2004. *Friends* featured six main characters—Rachel Green (Jennifer Aniston), Monica Geller (Courteney Cox), Phoebe Buffay (Lisa Kudrow), Joey Tribbiani (Matt LeBlanc), Chandler Bing (Matthew Perry), and Ross Geller (David Schwimmer), who are friends in their 20s and 30s and live in Manhattan.

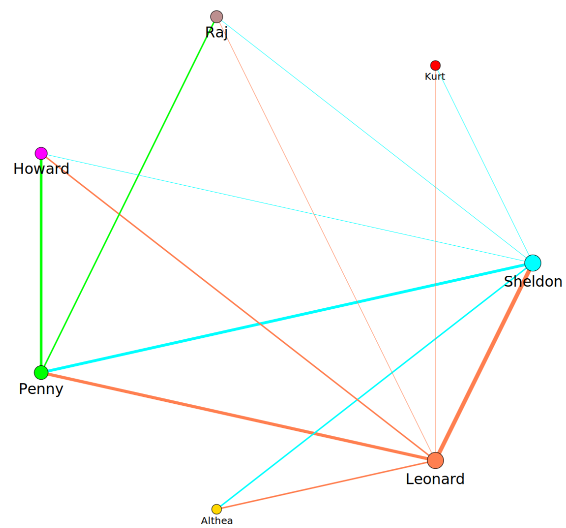

Fig. 6: *The Big Bang Theory*: Network of Characters (pilot).

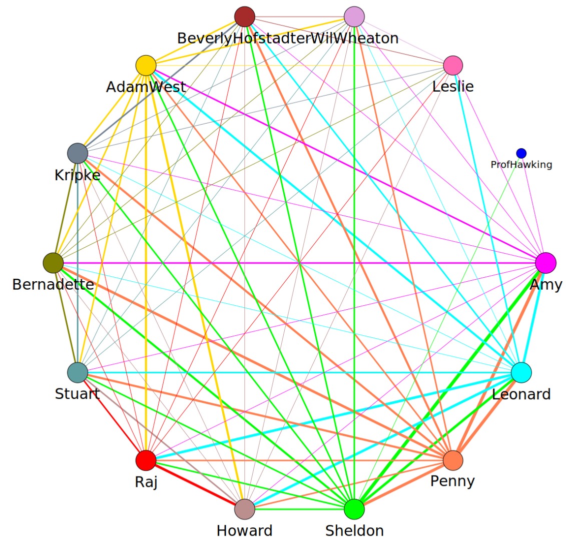

Fig. 7: *The Big Bang Theory*: Network of Characters (big clique).

The story unfolds at three main settings: a Manhattan coffeehouse (Central Perk) and the apartments across the hall, where the pairs of characters Monica and Rachel and Joey and Chandler live. *Friends* is about young people in a big city coming together to share living expenses, far from their parents, where friends act as surrogate family members.

Friends is frequently associated with these facts: (i) all six characters are equally prominent, i.e., they are generally given equal weight across the series; (ii) *Friends* is a multistory sitcom with no dominant storyline; (iii) *Friends* is the first true "ensemble" show—a series with no clear star or center, a cast of equals; (iv) the creators of *Friends* felt that six equal players, rather than emphasizing one or two, would allow for myriad story lines; (v) Monica likes to consider herself as hostess / mother hen; (vi) friendship as surrogate family; (vii) Ross and Rachel have an intermittent relationship.

Most of these facts have a key influence in the structure of the social network. For instance, the cliques normally involve all six characters, and the density of interactions is low.

Two social networks of *Friends* are shown in Fig. 4 and Fig. 5. These depict interactions spanning over all episodes in Season four, and interactions that happen in all Thanksgiving episodes respectively. In the former figure, all characters are depicted with same vertex size; in the latter, the size of the vertices is proportional to their degrees. Vertices in gray are not very important as they have low centrality. In both graphs, it is visible that most of the characters (i.e., all that are non-core characters) never get to meet each other. This fact thus leads to overall low density.

V. The Structure of the Social Network of *The Big Bang Theory*

The Big Bang Theory, created by Chuck Lorre and Bill Prady and aired from 2007 to 2019 by CBS, is also an American sitcom. The most prominent characteristic of *The Big Bang Theory* is that it revolves around science (in particular, physics), being centered on scientists living in Pasadena, the location of Caltech, where some of the characters interact.

In its first seasons, it featured five characters: the geeky and socially clueless physicists Sheldon Cooper (Jim Parsons) and Leonard Hofstadter (Johnny Galecki), who work at Caltech share an apartment; the waitress and aspiring actress Penny (Kaley Cuoco), living across the hall; Howard Wolowitz (Simon Helberg), an aerospace engineer; and astrophysicist Raj Koothrappali (Kunal Nayyar). Both Howard and Raj work at Caltech as well. From Season four, two other characters get more prominent roles: Bernadette Rostenkowski (Melissa Rauch), a microbiologist who dates and later marries Howard, and the neuroscientist Amy Farrah Fowler (Mayim Bialik), Sheldon's girlfriend.

This show has innovated by featuring most of the characters who, as mentioned, are young scholars with high IQ, having studied in famous universities in the U.S.

and with outstanding achievements in science. On the other hand, these characters have social issues like not grasping some social norms (Sheldon), not being able to talk to women (Raj), living with his mother (Howard), and resenting his mother's lack of attention toward him (Leonard). These issues often create innovative situations, with elaborate forms of communicating scientific facts. Besides, the more intellectual characters tend to interact more among themselves.

Again, as an illustration, Fig. 6 and Fig. 7 depict two social networks: the one corresponding to the very first episode, and the one with a big clique (Episode 17 in Season 9).

As shown in Fig. 6, where the node size represents the number of connection in that episode, both Raj and Howard were less prominent characters, when compared to Penny, Sheldon and Leonard. This is a probable consequence of the original idea of the show of making it about the relationship of the latter three (the original pilot, featuring a different actress for the role of Penny, was originally called, Sheldon, Leonard and Penny),

As the show developed, not only Raj and Howard became more central, but the two additional female leads were added to the central clique, resulting in the figure in the right, with the six characters having equally large nodes.

VI. Structural Characterization of the Social Networks

In this section, tools of network theory are used to characterize and facilitate the quantitative comparison of the networks of the aforementioned sitcoms. These tools and their respective measures are discussed briefly ahead. The reader is referred to these works for more details: [1], [2], [4], [6], [7], [5], [14], [25], [23], [24], [32]. Specifically for applications of these concepts on characterization of networks of TV series and other fiction works, please consult [8], [9], [10], [19]. These measures were mostly computed using *igraph* [15] for *python*.

While this section focuses on measures that refer to the networks as a whole (i.e., characterize the overall graph), the next section discusses measures that related to individual vertices or characters.

In order to be able to compare social network structures, data was collected manually by the author of the present paper, thus guaranteeing that the same criteria for extraction of relationships were used. Each episode was watched and notes were taken, regarding how many times who interacts with whom, i.e., based on the actual interactions of characters in each scene of each episode. An interaction happens when two characters talk (even if one talks and the other just listens) or touch or have eye contact. This means that, since not necessarily every character does interact with all others in a scene, each scene is not a complete graph. Thus, there are some differences between the way graphs are constructed using this method and, for instance, methods employed in [11], in [18], and in others, where automated techniques (e.g., language processing) are used to acquire data. The automated way generates complete graphs since all characters in a scene are connected to each other, even if they do not meet (for example, some leave the scene before others enter). As pointed out by [17], [16], this has an effect in some measures used to perform the analysis. While most of the metrics are not affected by automated data extraction, those related to clustering are not reliable in the automatically extracted networks.

The data used in the present paper refers to all seasons of *Seinfeld* and *Friends*, while it covers seasons 1–10 of *The Big Bang Theory*. This way, the data about the three sitcoms cover approximately 10 seasons. These seasons correspond to a different number of episodes, as shown in the corresponding column of Table I. Please notice that the number of episodes per season may vary according the source and/or due to two episodes being aired together; here data from imdb.com was used.

Networks are formalized as $G = (V, E)$, where V is the set of vertices in the graph, i.e., $|V|$ denotes the cardinality of the set of vertices. $E \subseteq V \times V$ is the set of edges (here an edge means a connection between two characters, i.e., an interaction in a scene); edges are weighted by the number of times two characters interact.

Some standard measures on G are:

- Density: ratio between the actual number of edges and the total number of possible edges;
- Geodesic shortest path between vertices i and j: path with minimum length that leads from i to j;
- Diameter: maximum geodesic path in the graph, i.e., maximum distance;
- Clique: subset of vertices of a graph, in which any two vertices are directly connected;
- Clique number: cardinality of the clique;
- Clustering coefficient: measures the ratio of connected triplets (3 vertices fully connected) by the ratio of possible triangles in a graph; it is a measure of how likely two neighbors of a vertex are connected.

Some of these are used to characterize the aforementioned networks. The results can be seen in Table I.

Columns titled $|V|$ and $|E|$ show how many vertices and edges respectively there are in each of the three

Table I: Graph's Characteristics of the Three Shows.

Graph	$\|V\|$	$\|E\|$	Density	Diam.	Clique Nb.	Clust. Coef.	Avg. Degree	Nb. of Episodes	Norm. $\|V\|$	Norm. $\|E\|$
Seinfeld	1448	9951	0.0095	6	11	0.0167	13.7	173	8.4	57
Friends	746	16569	0.06	5	10	0.03	44.6	236	3.2	70
The Big Bang Theory	311	12065	0.25	5	12	0.091	77.6	231	1.35	52

networks. Since the number of episodes (column 9) differs, the last two columns of that table show the respective normalized quantities. It is possible to see that the normalization does not change the order, but only the magnitude of number of vertices and edges.

As typical for such shows, lots of characters never get to meet others. This can be seen by two of the measures shown in Table I. The column that regards the clustering coefficient tells us how likely two neighbors of a given vertex are connected, i.e., how likely it is that two friends of a character also interact with each other. The graph's clustering coefficient measures the ratio of connected triplets (3 vertices fully connected) by the ratio of possible triangles. Notice the relatively low clustering coefficients, which reinforces that some characters never meet. Of course, for individual episodes, clustering coefficients change a lot, and some values are high. Also density (column 4) tells us that the interactions are not dense, i.e., not everybody interacts with a lot of other characters. The network of *The Big Bang Theory* is surprisingly dense though. This is probably due to the low number of vertices. Indeed, the number of characters other than the core ones is much lower than in the other two shows. Further, a high number of scenes involve many of the core characters, with just few non-core ones.

The other two measures shown in Table I are the diameter and the clique number. Both do not differ much from sitcom to sitcom. The low diameter is due to the fact that virtually all of the characters do interact with someone in the core of characters, thus the degree of separation between any two vertices is low. As for the clique, it is never the case that more than roughly 10 characters all interact with all others in the clique. These cliques normally occur in a situation in which the core characters are present, plus a couple of others, such as a party, a dinner, or some sort of meeting.

A second characterization of these networks is through its change along time. Next, for each sitcom, some of the aforementioned metrics are shown, this time broken down by season.

Table II: *Seinfeld*: Graph's characteristics in different situations

Season	Episodes	$\|V\|$	$\|E\|$	Diameter	Clustering Coef.
s1	5	34	136	4	0.24
s2	12	83	537	4	0.20
s3	23	169	988	6	0.095
s4	23	220	1299	5	0.11
s5	21	202	1347	6	0.10
s6	23	217	1227	5	0.068
s7	22	224	1408	5	0.12
s8	22	263	1491	5	0.078
s9	22	221	1518	5	0.098

A. *The Networks of Seinfeld*

Table II depicts the number of vertices and edges, plus diameter and clustering coefficient for each season of *Seinfeld*. Please notice that the number of episodes varies a lot, as mentioned in Section III, thus the column titled "Episodes" highlights this fact.

The diameter does not change much, being around 5 in each season. The clustering coefficient changes from season to season, as the cardinality of V and E changes.

As widely known, these three sitcoms were very popular among viewers. More recently, viewers have added ratings to each episode. Hence, a question arises whether such ratings correlate with any given structural measure of the corresponding network. The ratings per episode for *Seinfeld* are shown in Fig. 9. These values stem from imdb.com.

As mentioned in Section II, authors in [13] have shown that, for *The Big Bang Theory*, there is a high correlation between the decline in the concentration of dialogues and a decline in ratings of the show. Whereas the present work does not focus on the actual dialogues, an alternative way to measure concentration around some characters, per episode, is to measure how the degrees are distributed (or, rather, concentrated).

In order to represent such distribution of degrees it is useful to measure the entropy of such distribution.

This measure is based on orbits or partitions formed by the distribution of degrees. Denoting as n_l the cardinality of a subset l, and, further, assuming that G is partitioned into k subsets in which in each subset l all vertices have the same degree, the entropy H of a graph G is given by Eq. 1. Note that each episode corresponds

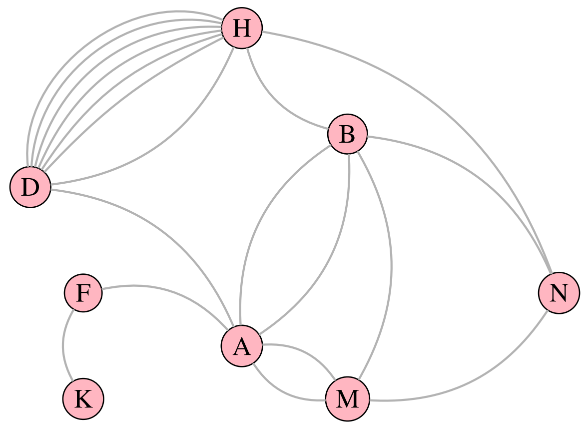

Fig. 8: An Example Graph with 8 vertices.

to one graph G.

$$H(G) = -\sum_{l=1}^{k} \frac{|n_l|}{n} * log \frac{|n_l|}{n} \quad (1)$$

In order to compare graphs of different sizes, it is useful to normalize the entropy so that entropy ranges between zero and one.

An example graph, partitioned according to the distribution of degrees, with normalized entropy close to zero would be a star. For instance, take a star with 1000 vertices. One of them has degree 999. All the others have degree 1. Thus, using Eq. 1 would yield a value close to zero. On the other hand, a graph with entropy 1 is one in which each vertex has a different degree. An example is shown in Fig. 8, where the entropy is 1 given that there are eight subgraphs, each containing one vertex since all vertices have different degrees.

An statistical analysis of the degree distributions of each episode of *Seinfeld* and the rating values of the corresponding episode show that the correlation rate is low, thus one cannot say that those two quantities are correlated for this show.

B. *The Networks of Friends*

A similar analysis was performed for the sitcom *Friends*; results are in Table III. In this sitcom, all seasons have 24 episodes, except the last one, which has only 18. Hence these values are not shown in the table. Again, note that some media may have merged two episodes.

Table III shows pattern similar to *Seinfeld* for diameter, while the clustering coefficient is more constant in *Friends* than it is in *Seinfeld*.

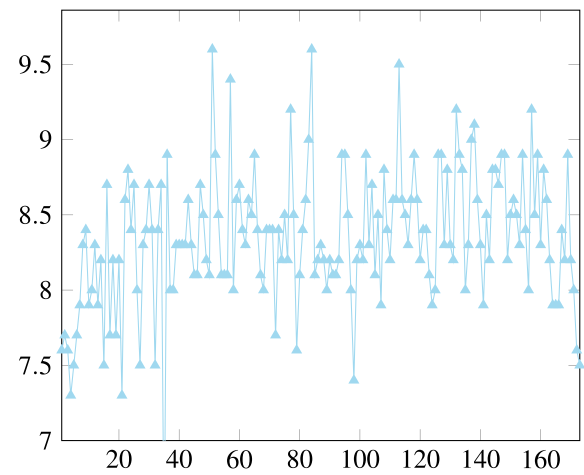

Fig. 9: *Seinfeld*: Ratings through episodes.

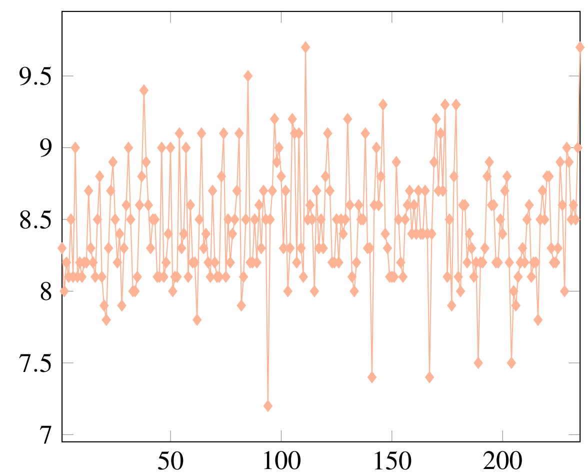

Fig. 10: *Friends*: Ratings through episodes.

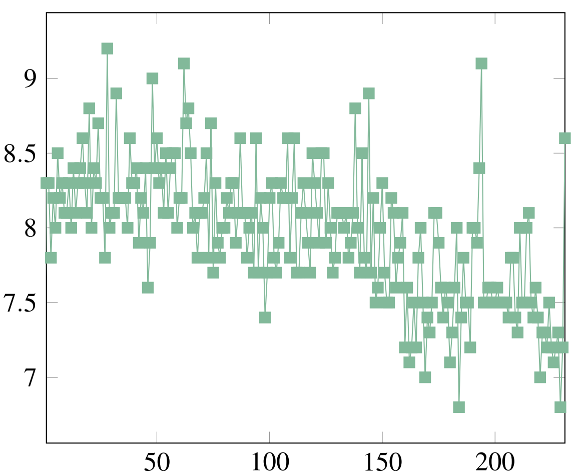

Fig. 11: *The Big Bang Theory*: Ratings through episodes (seasons 1 to 10).

Table III: *Friends*: Graph's characteristics in different situations

Season	N	$\|E\|$	Diameter	Clustering Coef.
s1	126	2492	5	0.16
s2	107	1815	5	0.19
s3	98	1770	5	0.20
s4	96	1598	4	0.23
s5	93	1786	4	0.19
s6	99	1491	4	0.16
s7	81	1475	5	0.20
s8	110	1220	4	0.14
s9	101	1454	4	0.19
s10	87	1468	5	0.23

Table IV: *The Big Bang Theory*: Graph's characteristics in different situations

Season	N	$\|E\|$	Diameter	Clustering Coef.
s1	50	799	4	0.25
s2	48	1020	3	0.27
s3	56	922	3	0.22
s4	50	1187	4	0.31
s5	47	1019	3	0.30
s6	50	1058	5	0.35
s7	55	1487	4	0.37
s8	48	1351	4	0.39
s9	45	1717	3	0.49
s10	39	1505	3	0.43

The ratings per episode of *Friends* appear in Fig. 10. Here too a low correlation was observed between such ratings and the entropy of the distribution of degrees.

C. The Networks of The Big Bang Theory

Finally, Table IV shows the values for the metrics applied to each season of the sitcom *The Big Bang Theory*. All seasons of *The Big Bang Theory* have either 23 or 24 episodes, except for the first season, which has 17. Note that the diameters are lower than for the seasons of *Seinfeld* and *Friends*, whereas the clustering coefficients are higher. This is in line with the conclusions reported in the beginning of the current section, which is related to the whole graph, rather than broken down by season.

Regarding popularity, the ratings of *The Big Bang Theory* went down with the advance of the seasons, as shown in Fig. 11. This may be a reason why [13] decided to investigate whether this could be explained by some measure. However, while these authors have observed a correlation between ratings and concentration of dialogues, no correlation between ratings and entropy of the degree distribution was observed.

VII. Characterizing the Vertices

The discussion presented in Section VI is useful to recognize patterns of interest in the sitcoms and also to compare them. However, for drawing conclusions (and perhaps establishing comparisons) on the individual characters of the sitcoms, one needs to use metrics relating to each vertex. Mostly, these refer to centrality measures. For more details on these, please refer to [8], [14].

Three metrics that characterize a given vertex of a graph in terms of its centrality are:

- Degree of vertex i: number of connections a vertex has, including multiple direct connections of i to other vertices;
- Betweenness b_i of vertex i: number of geodesic paths from vertex s to vertex t that pass through i; this quantity can be rescaled by dividing it by the number of pairs of vertices not including i so that $b_i \in [0,1]$;
- Closeness of vertex i: inverse of the mean shortest distance from i to each of the other vertices.

Given that the diameters are low, the values regarding closeness of vertices tend to be similar, no matter the importance of the character; henceforth these are not discussed.

Recall that each sitcom has a small number of core characters.

Of particular interest is the question of how those centrality measures change along season. To start with, Fig. 12 shows how the (absolute) values of degrees change for the core characters in *Seinfeld*. Obviously, since the first two seasons had a much lower number of episodes, this fact has to be taken into account. Therefore, it is useful to normalize the degrees by the number of vertices. This is shown in Fig. 13 in a more compact way.

One can see that, if normalized, the degree values from season to season are close; they vary between 1 and 3, being the highest in Season 2.

In *Friends*, as mentioned, one distinguishing characteristic is that the six core characters have nearly the same importance. While it has been shown that the degree is in fact close for them, both when all 236 episodes are considered [9], as well as for each of the 22 first episodes [10], this can be seen also for each season in figures 14 (absolute values) and 15 (normalized).

A similar investigation for ten seasons of *The Big Bang Theory* is shown in figures 16 and 17 in absolute and normalized values respectively.

Finally, it is also interesting to analyze how the centrality values change for the core characters, when

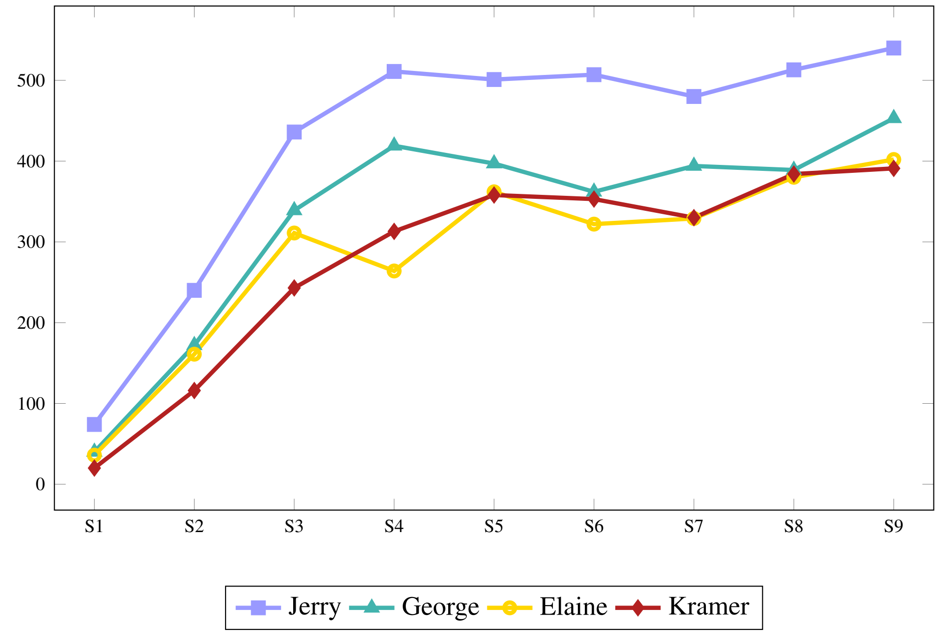

Fig. 12: *Seinfeld*: Degrees through seasons.

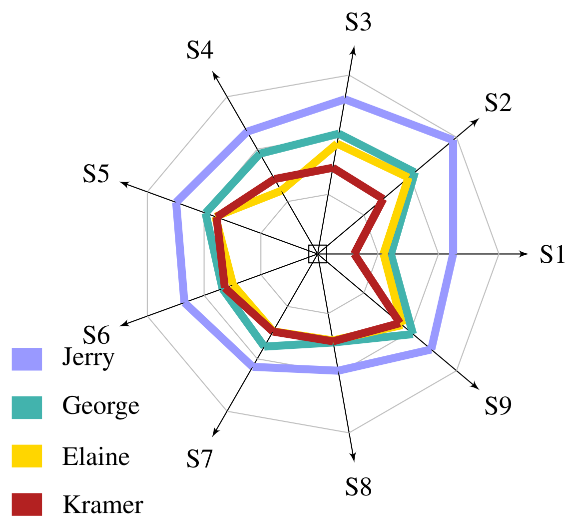

Fig. 13: *Seinfeld*: Normalized degrees through seasons.

compared within and across the three sitcoms.

Table V lists the values of degree and betweenness centrality (ordered by the latter, in each sitcom). Both centrality measures were normalized here by the size of the respective graph. One can see that in *Seinfeld* Jerry Seinfeld is in fact the character with the highest degree, but George, who has a degree close to Jerry's, has the highest betweenness (though also close to Jerry's). Then Elaine and Kramer follow, with similar centralities. These four characters have much higher centrality than the 5th character (Newman, Helen Seinfeld, and Morty Seinfeld appear closely in the 5th place).

The results for *Friends* were already mentioned, i.e., the six core characters have close degree. However, their betweenness are different. This means, as pointed out in [9], [8] that Joey is the guy who more efficiently connects a lot of other characters, while Monica, who has the highest degree, has low betweenness and thus she is not an impressive connector, being indeed the queen in her own apartment. Also worth noting is the huge difference from the six values to the 7th (a close tie between Judy Geller, Jack Geller, and Mike).

In *The Big Bang Theory*, as expected, Sheldon and Leonard are the most central characters, though Sheldon has a much higher betweenness centrality than all the others (even Penny's). Perhaps surprising, Penny has smaller degree than all other male characters, and her betweenness only surpasses Raj's. One explanation for this is that she has less social contact with characters other than the core ones. In fact, although she has a job, few scenes take place there and, when this happen, most of the core characters are there too. As for non-core characters, these have lower centrality when compared to the core ones. See for instance the last line of Table V, that relates to the 8th character (Stuart).

Also noteworthy is the fact that the degrees (even when normalized, as here) of the core characters are low in *Seinfeld*. The reason seems to be that this sitcom has

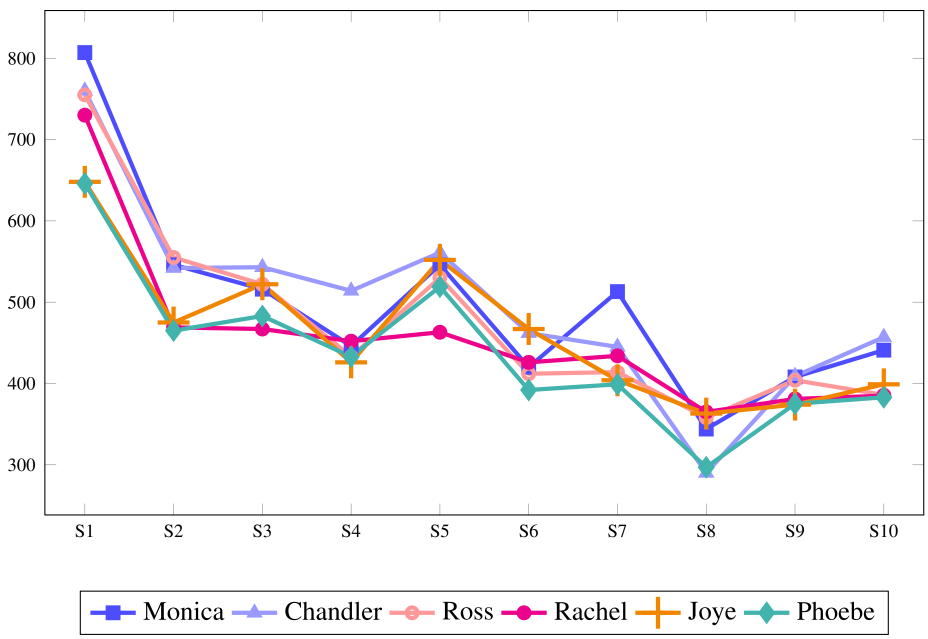

Fig. 14: *Friends*: Degrees through seasons

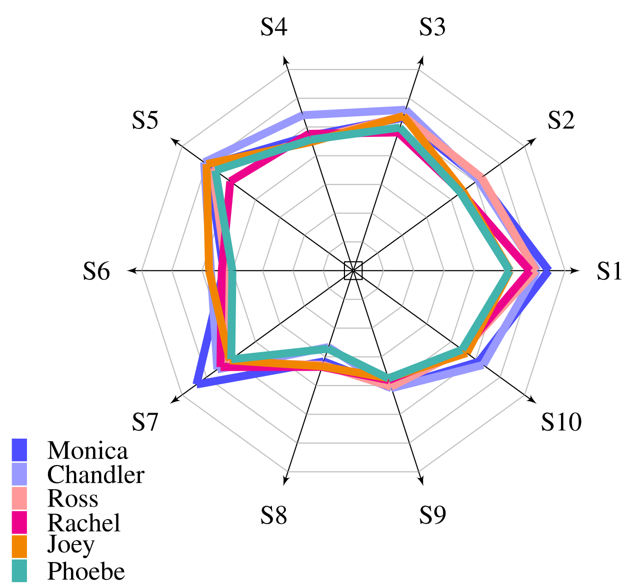

Fig. 15: *Friends*: Normalized degrees through seasons.

Table V: Betweenness centrality (alongside normalized degree) for all episodes.

Character	Norm. Degree	**Betweenness**
George	2.05	0.398
Jerry	2.63	0.387
Elaine	1.77	0.313
Kramer	1.73	0.286
5th	0.19	0.007
Joey	6.21	0.328
Ross	6.40	0.289
Chandler	6.69	0.240
Rachel	6.14	0.245
Phoebe	5.90	0.202
Monica	6.69	0.188
7th	0.24	0.005
Sheldon	13.54	0.489
Leonard	13.23	0.234
Howard	10.80	0.206
Penny	9.52	0.174
Raj	10.1	0.166
Bernadette	5.13	0.035
Amy	5.63	0.021
8th	1.84	0.013

the highest number of characters, as shown in Table I. Degrees in *Seinfeld* are roughly one third of those in *Friends*, while these are roughly half of those in *The Big Bang Theory*.

VIII. Concluding remarks

In this paper, three sitcoms that are centered on friendship among groups of young peole – *Seinfeld*, *Friends*, and *The Big Bang Theory* – were analyzed using techniques from network theory. The social networks of these sitcoms were formed by manually collecting pair-wise interactions in each scene. Although these networks

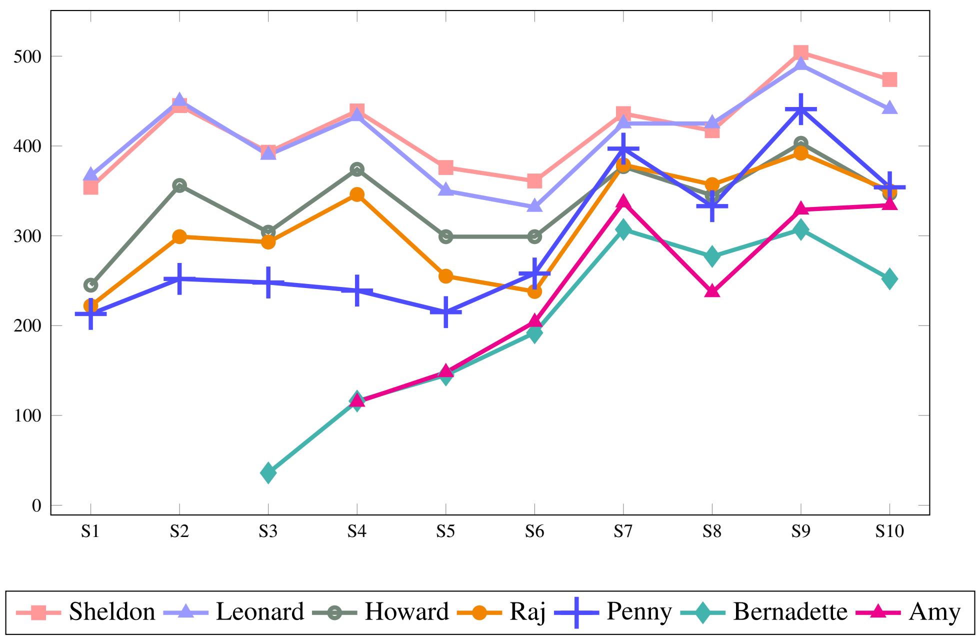

Fig. 16: *The Big Bang Theory*: Degrees through seasons

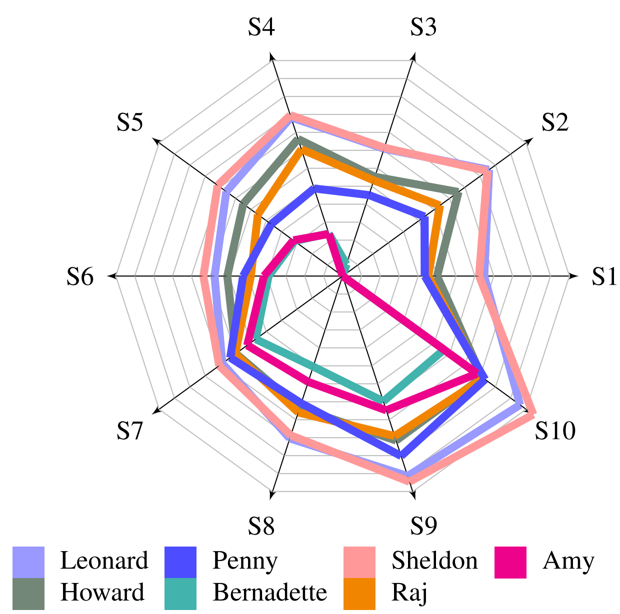

Fig. 17: *The Big Bang Theory*: Normalized degrees (seasons 1 to 10) for the seven main characters; max. value is 12.5 for Sheldon in Season 10

do not consider length of scene or even emotionally charged interactions, they can be a reasonable indicator of some aspects of the shows.

By doing the math related to the respective social network, it is possible to come to the following findings. The number of characters varies a lot throughout the shows, as do the number of interactions. This causes some networks to be more dense than others. The average degree of the characters also varies, not only among the core characters (with the exception of *Friends*, a show that is in fact well known for this feature), but also between the shows. While characters in *Seinfeld* have low average degree, those in *The Big Bang Theory* have high degrees. Measuring centrality by betweenness shows some surprises since those characters who are expected to be less central do in fact have high betweenness, thus acting as connectors between characters.

Regarding the temporal aspect, it was shown that graphs for different time slices (seasons or episodes) of the show change. It remains to be investigated which is the role of the techniques used by [26], where specific centrality measures for temporal networks were used.

This paper has also analyzed whether it is possible to credit the change in rates of each episode of the shows by the variability in the degrees (measured by the entropy), concluding that it was not possible to state that there is such a correlation.

There might be other ways to analyze the variability in the rates, which is left as future work.

ACKNOWLEDGEMENTS

Ana Bazzan is partially supported by CNPq (grant no. 307215/2017-2). This work was partially supported by

CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil) - Finance Code 001.

Referências

[1] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74(1):47–97, Jan 2002.

[2] Réka Albert, Hawoong Jeong, and Albert-László Barabási. Internet: Diameter of the world-wide web. *Nature*, 401(6749):130–131, 1999.

[3] A. Albright. The one with all the quantifiable friendships. https://thelittledataset.com/2015/01/20/ the-one-with-all-the-quantifiable-friendships/, 2015. (Accessed August, 2019).

[4] A. L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[5] A. L. Barabási, N. Gulbahce, and J. Loscalzo. Network medicine: a network-based approach to human disease. *Nature Reviews Genetics*, 12(1):56–68, January 2011.

[6] A. L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, and T. Vicsek. Evolution of the social network of scientific collaborations. *Physica A*, 311(3-4):590–614, 2002.

[7] Albert-László L Barabási and Eric Bonabeau. Scale-free networks. *Scientific American*, 288(5):60–69, 2003.

[8] A. L. C. Bazzan. I will be there for you: six friends in a clique, April 2018. ArXiv.

[9] Ana L. C. Bazzan. I will be there for you: clique, character centrality, and community detection in *Friends*. *Computational and Applied Mathematics*, 39(192), 2020.

[10] Ana L.C. Bazzan. Similar yet different: the structure of social networks of characters in *Seinfeld*, *Friends*, *How I met Your Mother*, and *The Big Bang Theory*. *RITA – Revista de Informática Teórica e Aplicada*, 27(4):66–80, 2020.

[11] Andrew Beveridge and Jie Shan. Network of thrones. *Math Horizons*, 23(4):18–22, 2016.

[12] Xavier Bost, Vincent Labatut, Serigne Gueye, and Georges Linarès. *Extraction and Analysis of Dynamic Conversational Networks from TV Series*, pages 55–84. Springer International Publishing, 2018.

[13] Andrea Fronzetti Colladon and Maurizio Naldi. Distinctiveness centrality in social networks. *PloS one*, 15(5):e0233276, 2020.

[14] Luciano da F. Costa, Francisco A. Rodrigues, Gonzalo Travieso, and P. R. Villas Boas. Characterization of complex networks: A survey of measurements. *Advances in Physics*, 56(1):167–242, 2007.

[15] Gabor Csardi and Tamas Nepusz. The igraph Software Package for Complex Network Research. *InterJournal*, Complex Systems:1695, 2006.

[16] M. Edwards, J. Tuke, M. Roughan, and L. Mitchell. The one comparing narrative social network extraction techniques. In *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 905–913, 2020.

[17] Michelle Edwards, Lewis Mitchell, Jonathan Tuke, and Matthew Roughan. The one comparing narrative social network extraction techniques. 2018.

[18] Milan Jasonov. Network science predicts who dies next in Game of Thrones. https://cns.ceu.edu/article/2017-07-08/network-science-predicts-who-dies-next-game-thrones, 2017. (accessed Jan. 9, 2018).

[19] Vincent Labatut and Xavier Bost. Extraction and analysis of fictional character networks: A survey. *ACM Comput. Surv.*, 52(5):89:1–89:40, September 2019.

[20] D. Liu and L. Albergante. Balance of thrones: a network study on 'game of thrones'. https://arxiv.org/abs/1707.05213, 2017. (Accessed August, 2019).

[21] J. Lv, B. Wu, L. Zhou, and H. Wang. StoryRoleNet: Social network construction of role relationship in video. *IEEE Access*, 6:25958–25969, 2018.

[22] Chang-Jun Nan, Kyung-Min Kim, and Byoung-Tak Zhang. Social network analysis of TV drama characters via deep concept hierarchies. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*, ASONAM '15, pages 831–836, New York, NY, USA, 2015. ACM.

[23] M. E. J. Newman. Scientific collaboration networks. I. network construction and fundamental results. *Phys. Rev. E*, 64:2001, 2001.

[24] M. E. J. Newman. Scientific collaboration networks. II. shortest paths, weighted networks, and centrality. *Phys. Rev. E*, 64:2001, 2001.

[25] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45:167–256, 2003.

[26] Sandra D. Prado, Silvio R. Dahmen, Ana L.C. Bazzan, Padraig Mac Carron, and Ralph Kenna. Temporal analysis of literary texts. *Advances in Complex Systems*, 19(03):1–19, 2016.

[27] Y. Seth. Who was the lead character in Friends? the data science answer. https://yashuseth.blog/2017/12/29/data-analysis-lead-character-of-friends-data-science/, 2017. (Acessed Dec. 30, 2018).

[28] G. Simchoni. The one with friends. http://giorasimchoni.com/2017/06/04/2017-06-04-the-one-with-friends/, 2017. (Accessed August, 2019).

[29] J. Stavanja and M. Klemen. Predicting kills in game of thrones using network properties. https://arxiv.org/abs/1906.09468, 2019.

[30] S. Stoltzman. Seinfeld characters – a post about nothing. https://www.stoltzmaniac.com/ seinfeld-characters-a-post-about-nothing/, 2016. (Accessed August, 2019).

[31] M. S. A. Tan, E. A. Ujum, and K. Ratnavelu. A character network study of two Sci-Fi TV series. *AIP Conference Proceedings*, 1588(1):246–251, 2014.

[32] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):397–498, June 1998.