Analyzing Social Interaction between "SpongeBob SquarePants" Television Series Characters

Liad Nahmias
Information Systems Engineering
Ben-Gurion University of the Negev
Israel
liadn@post.bgu.ac.il

Michael Khavkin
Information Systems Engineering
Ben-Gurion University of the Negev
Israel
khavkin@post.bgu.ac.il

Abstract

Extracting social networks from literature has recently become a well-studied problem, focusing on deriving interesting patterns from such text. Literary text is not only limited to books or movies, but can also include television shows and series. The main task we explore in this project is detecting the special properties of social dynamics in the specific domain of television show transcripts, using various NLP (Natural Language Processing) tools. Particularly, we analyze the interaction between the main characters in the "SpongeBob SquarePants" animated television series. We focus on two types of analysis: extraction of topics in the conversation of the main characters, and distinguishing the sentiment of the conversations between the characters. Finally, we place these parameters on a graph to visualize our results.

1 Introduction

Studying and analyzing literary text has recently become one of the common tasks in the field of Natural Language Processing (NLP). The major interest in such tasks can be mainly attributed to the availability of data on the Internet and other mass media sources, such as television and other video services.

Analysis of literary text is not only limited to books and other written literature, but can also be conducted on Internet blogs, transcript of movies and television shows. The challenge in obtaining verbal data is related to the fact that this data is not always publicly available. To overcome such challenges, crawlers are implemented into many websites containing the text transcripts of the movies and shows, and extract this unstructured data.

The applications of analyzing literary data are not only academic, but also relate to other industrial purposes. For example, analysis of posts and blogs published online may reveal malicious activities of individuals, by considering the interaction between them. Other, more obvious applications, are increasing the popularity and exposure of television shows, by studying the most affecting interactions and enhancing them. Above all, such types of analysis enable us to reveal many hidden patterns in the text, and explore new knowledge about its subjects.

As an example of a television series, we focus on the transcripts of the **"SpongeBob SquarePants"** animated television series, which is an American animated television series created by marine biologist and animator Stephen Hillenburg. First broadcasted in 1999, it presents diverse social interactions and dynamics between its characters, thus easy to study in research. The main character of the show is SpongeBob, who triggers almost all the interactions in the series, and hence we consider the roles of the main characters in the series from his perspective. Figure 1 summarizes the main roles of the main characters in the series from *SpongeBob*'s perspective.

The "SpongeBob SquarePants" animated television series consists of more than 11 seasons and more than 229 different episodes. Each episode is divided into two 15 minutes parts. We extract the transcripts of 203 episodes, resulting in 406 transcripts. We then parse them, so that it is clear who is the speaking character and to whom its is speaking. We use the generated transcript files to apply a LDA (Latent Dirichlet Allocation) topic modeling on the conversations of the characters to study their social background in the show. Then, we classify the conversations of each character to their sentiment, performing sentiment analysis.

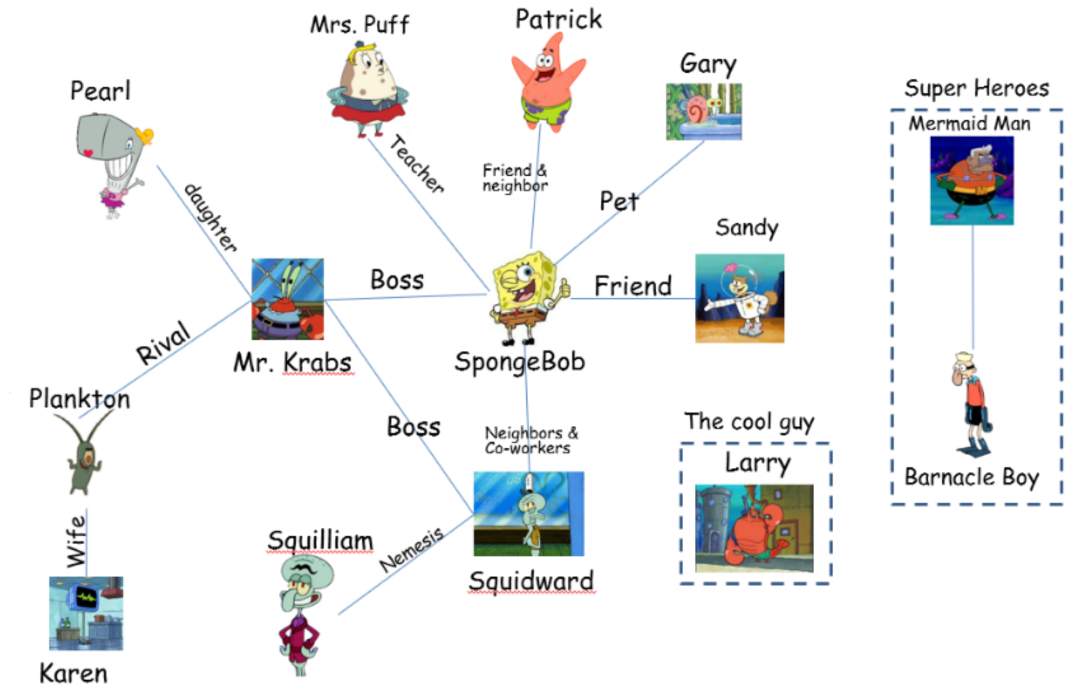

Figure 1: Roles and relationships between the main characters

We compare the results of these two models with our knowledge of the show, placing ourselves as domain experts. Our main goals are to evaluate the correlation between an expert analysis and an algorithmic NLP analysis, and determine whether characters with positive sentiment effect negative characters in their conversations. In addition, we group the characters to groups (i.e., communities) based on the main topics of their conversations.

In the following sections, we briefly list some works dealing with the analysis of literary text. We then describe the data used in our evaluation. We present our main hypotheses and explain the methodology used in our experiments. Finally, we present the results of the evaluation, and conclude with directions for future work.

2 Related Work

Analysis of text, in general, and literary text, in particular, is achieved using various methods. The analysis can be performed at the lexical or semantic levels, depending on the origin of the analyzed text and the domain in which it is used.

The task of understanding the ways in which information achieves widespread public awareness, using transcripts of famous movies was recently studied (Danescu-Niculescu-Mizil et al., 2012). The effect of phrasing on the memorability of movie quotes was studied using both a human test and a language model, built according to corpus of newswire. A controlled corpus of movie quotes was built to compare the properties of memorable and non-memorable quotes.

Another recent research studied the interactions between characters in the nineteenth-century British novels and serials (Elson et al., 2010). A method for extracting social networks from literature has been proposed. The construction of the social graph depends on the ability to determine when two characters are in conversation. Features from the social networks have been extracted to examine their correlation with one another, as well as with the metadata of the text. Interesting explanations for the differences in the social dynamics of the text are provided as an alternative to the explanations of literacy scholars.

3 Data

We extract the transcripts of 203 episodes, using a crawler we have implemented for the website `http://spongebob.wikia.com/wiki/List_of_transcripts`. Each episode is divided into two parts, resulting in 406 transcript files.

3.1 Pre-processing

Since the raw transcript does not identify the character to whom the text is addressed, we have manually added the "To" element for each processed conversation. Each transcript is saved in a structured format of <From>-<To>: <Text>, to easily pre-process the text.

In case a sentence was spoken from one character to two or more characters, we copied the spoken conversation as a new conversation for each receiving character. For example, if *SpongeBob* addresses *Patrick* and *Squidward* with the line "Good Morning Everyone", we have created two separate conversations of the form: <*SpongeBob*>–<*Patrick*>: "Good Morning Everyone" and <*SpongeBob*>–<*Squidward*>: "Good Morning Everyone". Figure 2 shows an example of the extracted transcripts.

For the analysis, we have only focused on 14 main characters: *SpongeBob, Patrik, Gary, Squidward, Sandy, Mr. Krabs, Mrs. Puff, Squilliam, Larry, Pearl, Plankton, Karen, Mermaid Man* and *Barnacle boy*. We tried to be diverse enough in the selection of different episodes, to generate data that covers different themes and scenarios.

In addition, we have created two copies of the data. In the first, we tokenized each conversation and removed stopwords and any punctuation sings. This dataset was used for the topic modeling. The second copy contained the raw text, containing stopwords and other punctuation signs. We used the second dataset for the sentiment analysis, to increase its accuracy, since punctuation marks (such as the exclamation mark) can imply sentiment.

4 Hypotheses

It is clear that each episode in the series deals with a different story, and thus the interactions may differ between episodes. Nonetheless, all episodes share the same social background, and thus we hypothesize that the topics, extracted from a character's conversations are not different between episodes, and that the episodes we have collected are representative enough of the whole series. In addition, we perform a comparison between an expert analysis and an algorithmic NLP analysis, both in terms of the sentiment score and the expected topics, to check whether there is some correlation between a human analysis to that conducted by a computerized algorithm.

Focusing on the sentiment analysis, we consider the average sentiment score of all the conversations spoken by each character, and compare it to the sentiment of each such conversation separately. We hypothesize that character with positive average sentiment effects a negative character in their conversations. While this is difficult to assess, we focus only on conversations spoken by the six key characters, where the conversations are long enough to produce accurate sentiment score.

Lastly, we hypothesize that the topics extracted from a character's conversation depend on his dynamics with other characters. Similarly to the comparison above, we compare the extracted topics from all the conversations of the main characters to the topic labels we manually set as a human annotation.

5 Social Interaction analysis

In this section, we describe in detail the algorithmic methods used in our evaluation. All algorithms are executed within Python 2.7 on a Intel Core i7 2.5 GHz computer with 16 GB RAM and Windows 10 OS.

5.1 Modeling topics in interactions

We applied the LDA (Latent Dirichlet Allocation) algorithm, implemented in the Gensim python package, to extract the main 50 topics from the all our corpus (containing the conversations between all the main characters). We then, locate the most coherent topics and match them with the different characters. We label each topic relating to each character using the top 15 terms in the topic.

In addition, we used a visualization tool, called pyLDAvis, to visualize the topics and extract the terms, for which the overall term frequency equals to the estimated term frequency within that topic. The visualization was set to display the top 30 most relevant terms of that topic. We consider the first ten terms to indicate the overall social status of the character.

From	To	Text
Squidward	Squidward	Oh what a beautiful day. And here I am trapped in a prison of high cholesterol.
BackGround	BackGround	bell dings
Squidward	Squidward	No one ever comes in on Sunday.
BackGround	BackGround	bell dings again
Squidward	Squidward	Why can't Mr. Krabs just let us go home?
BackGround	BackGround	bell dings again. Squidward gets angry. The scene changes to to SpongeBob ringing a bell set on the order window. Squidward runs up to SpongeBob
Squidward	Spongebob	SpongeBob stop ringing this bell!
BackGround	BackGround	He picks it up and slams it on the bottom of the order window
Spongebob	Squidward	I was just testing it.
BackGround	BackGround	Squidward leans through the order window getting in SpongeBob's face
Squidward	Spongebob	I will ring the bell when there's an order. But
BackGround	BackGround	scene zooms out to show that the restaurant is empty
Squidward	Spongebob	there's no customers! There hasn't been one all day and there isn't gonna be any!
BackGround	BackGround	he struggles to pick up the cash register but successfully does so and he slams it down making a bell noise
Spongebob	Squidward	One Krabby Patty coming up!
Squidward	Spongebob	No!
BackGround	BackGround	register drawer shoots open knocking Squidward out of the way. A bunch of coins fall onto the floor. The scene changes to show Mr. Krabs office where I
Mr. Krabs	Mr. Krabs	That sounds like me money dropping.
BackGround	BackGround	The scene changes to show the outside of the office and Squidward is picking up the coins. Mr. Krabs opens his office door.
Mr. Krabs	Mr. Krabs	What's going on out here?!
Mr. Krabs	Money	My babies!
BackGround	BackGround	runs up to Squidward and shoves him away
Mr. Krabs	Squidward	Get away you barbarian! What have you done? Nice clean money soiled!
BackGround	BackGround	scoops up the coins in his claws
Mr. Krabs	Money	I'll take care of ya. Let papa clean ya up.
BackGround	BackGround	The scene changes to show a long shot of SpongeBob staring out from the order window
Mr. Krabs	Squidward	Clear the way!
BackGround	BackGround	he runs into the kitchen and starts washing them off in the sink
Mr. Krabs	Money	No no no don't cry little ones.

Figure 2: Format of file generated from the collected data

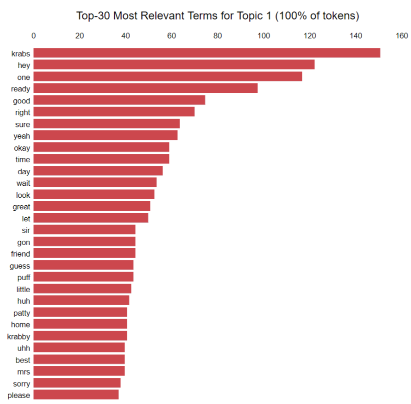

Figure 3: Topic model visualization for conversations of Spongebob, generated by the PyLDAvis package. The red bars indicate the estimated term frequency within the selected topic

5.2 Modeling sentiment in interactions

Since the relationships between the different characters stay the same across different episodes of the series, we classified each conversation to its sentiment (positive, negative or neutral), using the NLTK Python package. No stopwords or punctuation signs have been removed, to increase the accuracy of classification.

Then, as a gold standard, we used our knowledge (as domain experts for this task) to compare the performance of the classification, positioning the sentiment score in the range between -1 (strong negative sentiment) and 1 (strong positive sentiment), relatively to a threshold of 0.05 (fig. 4). We average the sentiment score for all the conversations of a character to get a sense of its general mood.

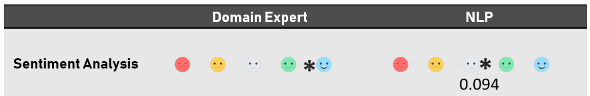

Figure 4: *SpongeBob*'s average sentiment score from his conversations, compared to score annotated by a domain expert. Red indicates strong negative sentiment, while blue indicates strong positive sentiment

5.3 Constructing the social interaction network

Following the idea behind the construction of the social network of Elson et al. (Elson et al., 2010), we used the "NetworkX" python package to construct the following directed graph:

1. **Vertices** – Each character entity was assigned to a vertex in the graph. We only included nodes corresponding to characters with a "PERSON" POS (Part of Speech) tag count of at least 2 (that is, all the characters who are mentioned more than two times in all the collected data). The *size* of each vertex was drawn to be proportional to the character's share in all the "PERSON" named entity mentions in the collected transcripts. The *color* of each vertex corresponds to common labels of its topics.

2. **Edges** – Each conversation (between two characters) is assigned a directed edge. The *direction* of the edge is used to indicate the overall sentiment of the sentences spoken by a character to another character. Green edges indicate positive sentiment, blue indicates neutral sentiment and red indicates negative sentiment. The *weight* on the edges shows the average score of the sentiment of all the sentences a character spoke during his conversation. The *width* of an edge is proportional to the length of the conversation, which is considered as the number of sentences that the character spoke, normalized to the total number of collected sentences.

6 Evaluation Results

In this section, we present the evaluation results. An overview of the social interaction graph is provided.

6.1 Sentiment Analysis

Table 5 compares the sentiment analysis of each main character, evaluated by us (as domain experts) and by the NLTK sentiment analyzer. Most of the results match our expectations (with minor deviations), except for *Plankton*, *Squidward* and *Mrs. Puff*, where almost opposite sentiment was detected. We attribute these changes to the varying conversation length and different text length for these characters.

We then compared the sentiment of each character with the sentiment dictated by his conversations with other characters, to check whether a character with positive sentiment effects a negative character in their conversations. The results of such comparison, regarding the six most dominant characters in the show are shown in table 6.

The differences between the sentiment scores of the characters and their conversations are not significant, implying that we can not significantly conclude that a positive character affects a negative character in their conversations (or vice versa).

One exception that we observed is the relationship between *Squilliam* and *Squidward*, which are designed to be rivals. Squidward, which tends to be neutral in his overall average sentiment (score of 0.052), turns his conversations with *Squilliam* to be more positive than *Squidward*'s average sentiment (score of 0.23, which is closer to positivity). But, since the results are not statistically significant, we cannot infer this for the entire set of characters.

There was little correlation between the sentiment analysis and the expert analysis. It seems that our sentiment score was most of the time close to the negative or positive polar of the scale, while the NLP sentiment score was mostly close to the threshold of neutrality. This can relate to the inefficient implementation of the sentiment analysis algorithm in the NLTK package.

Moreover, we have not found strong evidence that will imply that positive characters have an effect on negative characters. Furthermore, we can see that there is no effect on the sentiment between two characters.

6.2 Topic Modeling

Figure 3 shows the 30 most relevant terms of the topic, generated by LDA model, which matches the conversations of *SpongeBob*. For example, *SpongeBob* is employed by *Mr. Krabs* in his restaurant, and this corresponds to the most frequent term *"krabs"* in the LDA model. This is in line with our assumptions about the topics in his conversations.

Table 8 shows the results of topic modeling applied to the conversations of the main characters. Some of the topics are very coherent, while others are less interpretable. It is possible to find common topics between characters and compare them to their relationships. We group similar topics (with similar attached labels) of similar characters in the social network.

6.3 Social network

The constructed social network is shown in fig. 7. We arranged the nodes (corresponding to characters) in a circular form to increase readability. Recall that we place the most dominant character of *SpongeBob* in the center to create the social network relatively to him.

We can observe that most of *SpongeBob*'s conversations are dictated by a positive sentiment, and most of these conversations are directed to *Patrick*, *Sandy*, and his employer *Mr. Krabs*.

<table>
<tr><th></th><th colspan="2">Sentiment Analysis</th><th></th></tr>
<tr><th>Character</th><th>By domain expert</th><th>By analyzing the text</th><th>Average sentiment score</th></tr>
<tr><td>SpongeBob</td><td>[----------|------*--]</td><td>[----------|-*-------]</td><td>0.094</td></tr>
<tr><td>Patrick</td><td>[----------|-----*---]</td><td>[----------|---*-----]</td><td>0.057</td></tr>
<tr><td>Barnacle Boy</td><td>[--------*-|---------]</td><td>[---------*|---------]</td><td>-0.001</td></tr>
<tr><td>Mermaid Man</td><td>[----------|-*-------]</td><td>[-------*--|---------]</td><td>-0.043</td></tr>
<tr><td>Mr. Krabs</td><td>[----------|---*-----]</td><td>[----------|-*-------]</td><td>0.05</td></tr>
<tr><td>Larry</td><td>[----------|-----*---]</td><td>[----------|--*------]</td><td>0.035</td></tr>
<tr><td>Karen</td><td>[---------*|---------]</td><td>[----------|*--------]</td><td>0.02</td></tr>
<tr><td>Squilliam</td><td>[----------|-*-------]</td><td>[----------|-----*---]</td><td>0.36</td></tr>
<tr><td>Sandy</td><td>[----------|-----*---]</td><td>[----------|--*------]</td><td>0.053</td></tr>
<tr><td>Squidward</td><td>[--*-------|---------]</td><td>[----------|--*------]</td><td>0.052</td></tr>
<tr><td>Gary</td><td>[---------*---------]</td><td>[----------|*--------]</td><td>0.016</td></tr>
<tr><td>Pearl</td><td>[----------|--*------]</td><td>[----------|-*-------]</td><td>0.102</td></tr>
<tr><td>Plankton</td><td>[--*-------|---------]</td><td>[----------|--*------]</td><td>0.088</td></tr>
<tr><td>Mrs. Puff</td><td>[--------*-|---------]</td><td>[----------|--*------]</td><td>0.101</td></tr>
</table>

Figure 5: Average sentiment score for each character, compared to the score annotated by a domain expert

<table>
<tr><th colspan="3">Individual sentiment</th><th colspan="6">Conversational sentiment</th></tr>
<tr><th colspan="3"></th><th colspan="6">Conversation receiver ("To")</th></tr>
<tr><th></th><th>Character</th><th>Score</th><th>SpongeBob</th><th>Patrick</th><th>Mr. Krabs</th><th>Squilliam</th><th>Squidward</th><th>Mrs. Puff</th></tr>
<tr><th rowspan="6">Conversation speaker ("From")</th><td>SpongeBob</td><td>0.094</td><td></td><td>0.06</td><td>0.10</td><td>0.19</td><td>0.13</td><td>0.04</td></tr>
<tr><td>Patrick</td><td>0.057</td><td>0.05</td><td></td><td>0.07</td><td></td><td>0.14</td><td></td></tr>
<tr><td>Mr. Krabs</td><td>0.05</td><td>0.06</td><td>0.01</td><td></td><td></td><td>0.01</td><td>0.06</td></tr>
<tr><td>Squilliam</td><td>0.36</td><td>0.04</td><td></td><td></td><td></td><td>0.23</td><td></td></tr>
<tr><td>Squidward</td><td>0.052</td><td>0.04</td><td>0.08</td><td>0.05</td><td>0.08</td><td></td><td></td></tr>
<tr><td>Mrs. Puff</td><td>0.101</td><td>0.1</td><td></td><td>0.13</td><td></td><td></td><td></td></tr>
</table>

Figure 6: Comparison of the average sentiment of a character to the sentiment of his conversations with other characters. Gray cells indicate the absence of conversation between the corresponding characters

Character	Portion of spoken sentences
Background	41.92%
SpongeBob	25.22%
Squidward	9.49%
Patrick	7.16%
Mr. Krabs	6.04%

Figure 9: Portion of spoken sentences by the top 5 characters

Conversation	Portion of spoken sentences
Squidward<->SpongeBob	11.57%
Patrick<->SpongeBob	11.34%
SpongeBob<->Mr. Krabs	5.56%
SpongeBob<->Sandy	4.13%
Mr. Krabs<->Squidward	2.01%

Figure 10: Portion of spoken sentences by the top 5 conversations

Figure 11: Proportion of spoken sentences

In addition, we notice an overlap between the topics corresponding to the different characters. *Spongebob, Patrick, Squidward, Mr. Krabs* and *Squiliam* are grouped to a group (yellow) based on the common labels of their topics, relating to food and work. On the other hand, *Plankton, Mermaid-man, Barnacle Boy, Karen* and *Mrs. Puff* are grouped together (blue), since their topics mainly discuss actions of rivalry and competitiveness. The remaining characters of *Larry, Sandy, Pearl* and *Garry* discuss feelings and self-awareness and thus are grouped to a third group (Gray).

6.4 Additional Metrics

We measured the following features of the conversational network (as in Elson's et al. research (Elson et al., 2010)):

- Graph density, for the graph $G(V, E)$ where characters are represented with nodes V and their conversations with edges E. The density of the graph is $density = \frac{2*|E|}{|V|(|V|-1)} = 1/3$. This corresponds to the way we constructed our network, by considering conversations between two characters. In general, most of the dynamics in the series are characterized by conversations between 2–3 characters.
- The proportion of quotes spoken by each character are described in table 11. Most of the sentences were spoken by the narrator (Background), but as expected, SpongeBob

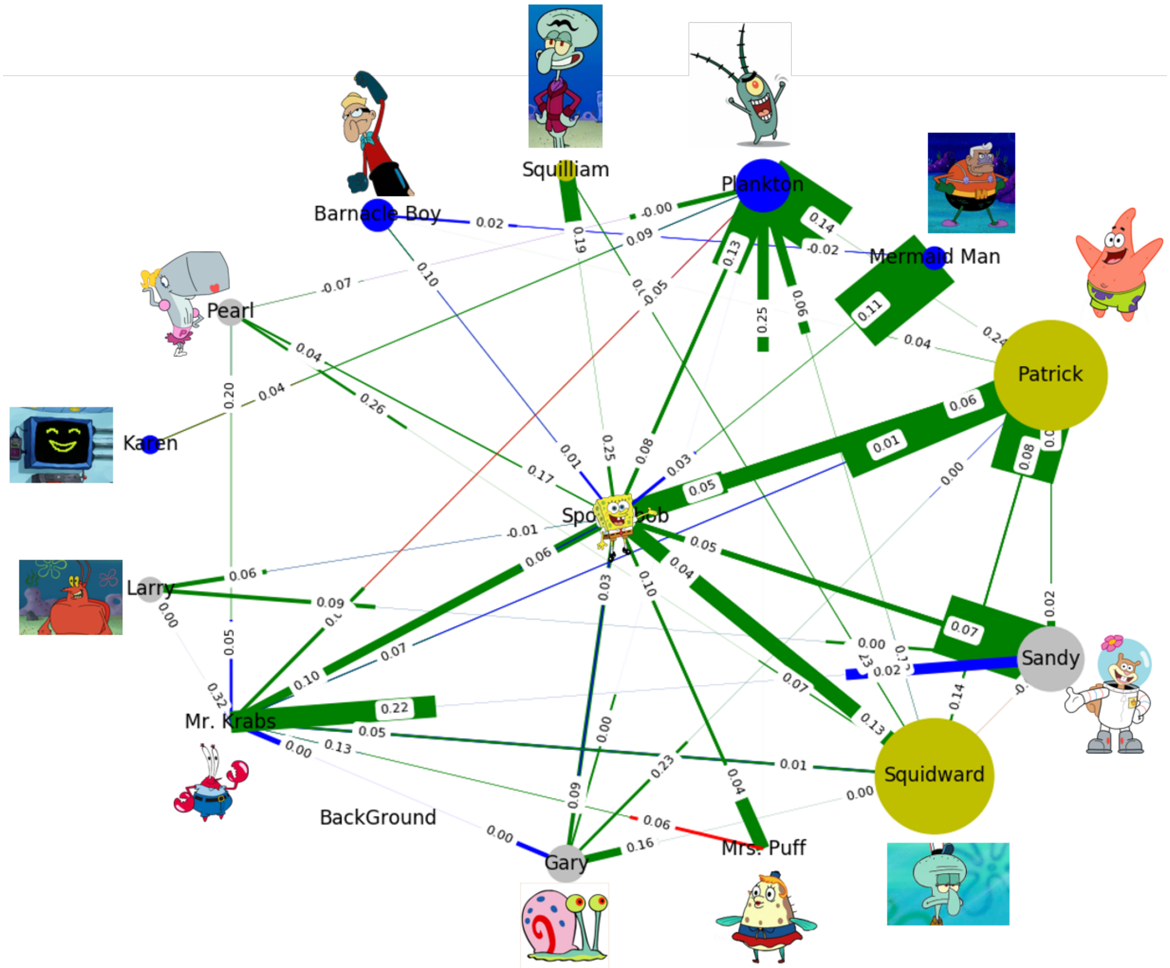

Figure 7: Directed graph representing the network of social interactions between the main characters of the show. The *weight* on the edges shows the average score of the sentiment of all the sentences a character spoke during his conversations. The *width* of an edge is proportional to the length of the conversation. The *size* of each vertex was drawn to be proportional to the character's share in all the "PERSON" named entity mentions. The *color* of each vertex corresponds to common labels of topics.

is the second most dominant speaker. Consequently, most of his conversations were with *Patrick* and *Squidward*. *Mr. Krabs* was also dominant, presenting the third highest portion of conversations with *SpongeBob*. We note that the text spoken by the background narrator is ignored in our analysis, since it is often noisy, but is interesting to explore as a future work.

7 Conclusions

In this project, we analyzed the social interactions of 14 main characters of the "SpongeBob SquarePants" animated television series. We have focused on two types of analysis using NLP tools: extraction of LDA model topics and a sentiment analysis of the conversations between the characters in the show. The expert analysis was performed by ourselves who are familiar with all the ten seasons of the show, while the NLP analysis was performed only on approximately two seasons of the show. However, we found some correlation between the results of each analysis. In addition, there was little correlation between the sentiment analysis and the expert analysis. It seems that our sentiment score was most of the time close to the negative or positive polar of the scale, while the NLP sentiment score was mostly close to the threshold of neutrality. We have not found strong evidence that will imply that positive characters have an effect on negative characters. Furthermore, we can see that there is no effect on the sentiment between two characters. In contrast, we can conclude that the dynamics between charac-

Character	15 Top words in each topic	Label
SpongeBob	krabs, hey, one, ready, good, sausages, hotter, sash-ringing, lima, yeh, drown, punch, national, anddone, vow	Work Commitments
Patrick	hey, good, gon, yeah, one, nets, puff, ate, dinga, living, tastes, bowl, club, borrowing, carni	Life style
Barnacle Boy	man, invisible, mermaid, come, boatmobile, let, signal, underwear, bob-acle, open, confound, jukebox, vision, bob-acle, show	Super heroes actions
Mermaid Man	evil, barnacle, boy, signal, man, call, waterball, rest, right, squash, yes, move, see, remember, mode	Super heroes actions
Mr. Krabs	money, boy, one, day, wait, bib, relay, shenanigan, suggestions, prized, photogeneric, years, okie, lass, scrape	Money
Larry	muscle, hey, serious, pop-up, piston, car, ambulance, guys, tonight, ball, thanks, slam, jerky, love, little	Feelings
Karen	name, andcute, overload, vincent, gon, decisions, hard, majesty, gorgeous, day, mishap, baba, tuesday, stumbled, reboot	Computer performance
Squilliam	tentacles, restaurant, name, everything, band, toilet, flawless, tuesday, admit, pass, top, tuesday, pass, garden, soda	Work
Sandy	gon, hi-yah, little, texas, hey, little, apologize, burn, sign, house, warmed, miss, competition, horns, cheeks	Feelings
Squidward	one, day, time, krabs, two, button, situations, ball, inspection, annual, clear, inspection, flavor, co-cashier, ball	Work day
Gary	meow, one, night, dream, peru, ages, think, stain, dickinson, peru, ages, think, stain, dickinson, peru	Thoughts
Pearl	new, daddy, krab, punch, krusty, pal, hate, kind, nothing, fry, nothing, hate, okay, uhh, sponginess	Feelings
Plankton	krabby, krabs, patty, one, puppy, stir, attention, nostrils, argh, snout, krabs, time, parry, krabby, puppy	Mr. Krab's restaurant
Mrs. Puff	class, thing, school, boat, first, date, months, proper, poseidon, old, congratulations, improvement, pedestrian, millions, excellent	Teaching

Figure 8: Topic modeling of the conversations of the main characters ($n_{topics} = 50$). Most coherent topics were matched to their corresponding characters and their annotated labeling

ters affect the topics of their conversations. Contrary to our hypothesis that each character has its own behavior and traits that keep evolving through the episodes, even many episodes that are more general and deviate from the main subjects of the show, do not affect the analysis of the characters, and this analysis is representative enough of the social dynamics. Directions for future work include extracting more transcripts to increase the size of the corpus and conducting other types of analysis concerning dynamics between characters, such as considering group conversations between more than two characters.

References

Cristian Danescu-Niculescu-Mizil, Justin Cheng, Jon Kleinberg, and Lillian Lee. 2012. You had me at hello: How phrasing affects memorability. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1*, pages 892–901. Association for Computational Linguistics.

David K Elson, Nicholas Dames, and Kathleen R McKeown. 2010. Extracting social networks from literary fiction. In *Proceedings of the 48th annual meeting of the association for computational linguistics*, pages 138–147. Association for Computational Linguistics.