

# The Movie Squid Game in BBC News: A Text Analysis

Author: Namkil Kang

Far East University

[somerville@hanmail.net](mailto:somerville@hanmail.net)

DOI:10.26821/IJSHRE.11.3.2023.110306

## ABSTRACT

*The ultimate goal of this paper is to analyze 27 pieces of BBC news regarding the movie Squid Game. The software package NetMiner was used to analyze 27 pieces of BBC news. To begin with, this paper provides information on the structural and descriptive properties of data analyzed by the software package NetMiner. To go into detail, network properties such as the number of isolated links, the number of a pendent, the number of links, their density, average degree, and the number of components have been accounted for. Also, we aim to examine network properties such as reciprocity (Arc and Dyad), transitivity, clustering coefficient, and a mean distance of links. A major point to note is that in the so-called word cloud, the noun Game is deemed to be the most significant, followed by the noun Squid, the noun show, and the noun Netflix, in that order. A further point to note is that topic 6 is the most occurred one in 27 pieces of BBC news, followed by topic 12, topic 7, and topic 9, in descending order. Additionally, this paper argues that the noun Game is the most widely used one in 27 pieces of BBC news, followed by the noun Squid, the noun show, and the noun Netflix, in that order. With respect to degree centrality, it is important to note that the noun Squid has the highest degree centrality. When it comes to closeness centrality, the noun Squid also has the highest closeness centrality.*

**Keywords:** squid game, BBC news, NetMiner, mining, topic, keyword, word cloud

## 1. INTRODUCTION

The main goal of this paper is to analyze 27 pieces of BBC news broadcasted from 11th, 2, 2020 to 2nd, 11, 2022 regarding the movie *Squid Game*. For this goal, the software package NetMiner was used to analyze 27 pieces of BBC news. First, we provide the structural properties and descriptive properties of data analyzed by the software package NetMiner. More specifically, we provide the number of isolated links and the number of a pendent. Also, we provide network properties such as the number of links, their density, average degree, and the number of components. We also provide network properties such as reciprocity (Arc and Dyad), transitivity, clustering coefficient, and a mean distance of links. Second, we provide a word cloud representing 27 pieces of BBC news in which major keywords occur differently in size. Third, we provide topic information on 12 topics and their keywords. Also, we look into how often major keywords occurred in the sentences constituting 27 pieces of BBC news. Fourth, we inquire into degree centrality that refers to the number of the linked neighbors. Also, we look into closeness centrality that indicates that the more the distance among nodes is close, the more they count as important or significant.

## 2. RESULTS

### 2.1. Structural and descriptive properties

In what follows, we aim at providing the structural and descriptive properties of data analyzed by the software package NetMiner. Table 1 shows information on the frequency of nouns and their

proportion and their cumulative proportion:

**Table 1 Frequency of nouns and their proportion**

Value	Frequency	Proportion	Cumulative Proportion
1.0	596	0.487	0.487
2.0	213	0.174	0.661
3.0	126	0.103	0.764
4.0	69	0.056	0.82
5.0	49	0.04	0.86
6.0	26	0.021	0.882
7.0	25	0.02	0.902
8.0	17	0.014	0.916
9.0	9	0.007	0.923
10.0	10	0.008	0.931
11.0	11	0.009	0.94
12.0	13	0.011	0.951
13.0	6	0.005	0.956
14.0	7	0.006	0.962
15.0	3	0.002	0.964
16.0	3	0.002	0.967
17.0	5	0.004	0.971
18.0	5	0.004	0.975
19.0	6	0.005	0.98
20.0	2	0.002	0.981

22.0	1	0.001	0.982
24.0	2	0.002	0.984
25.0	2	0.002	0.985
27.0	1	0.001	0.986
29.0	1	0.001	0.987
34.0	1	0.001	0.988
35.0	2	0.002	0.989
36.0	1	0.001	0.99
42.0	2	0.002	0.992
48.0	1	0.001	0.993
53.0	1	0.001	0.993
69.0	1	0.001	0.994
75.0	1	0.001	0.995
79.0	1	0.001	0.996
86.0	1	0.001	0.997
108.0	1	0.001	0.998
123.0	1	0.001	0.998
149.0	1	0.001	0.999
164.0	1	0.001	1
Total	1224	1	

It is interesting to note that one word appears in 27 pieces of BBC news and that its frequency is 596 tokens. More specifically, it has the highest frequency (596 tokens) and the highest proportion (0.487). Perhaps it is worthwhile pointing out, on the other hand, that there occurred two words whose frequency is 213 tokens (the second highest). Note that their proportion and cumulative

proportion are 0.714 and 0.661, respectively. It is particularly noteworthy that three words occur in 27 pieces of BBC news and that their frequency is 126 tokens (the third highest). Notice that their proportion and cumulative proportion are 0.103 and 0.764, respectively. It should also be pointed out that there occurred four words whose frequency is 69 tokens and that their proportion and cumulative proportion are 0.056 and 0.82, respectively. Now attention is paid to the descriptive properties of data analyzed by the software package NetMiner:

**Table 2 Descriptive properties**

Mean	Min	Max	Variance	Std. Dev.
3.92	1	164	99.801	9.99

As illustrated in Table 2, the mean (average) value of words occurred in 27 BBC news is 3.92 (words). Their minimum value is 1. This in turn indicates that one word has its minimum value. Note, on the other hand, that the maximum value of words occurred in 27 pieces of BBC news is 164. This in turn suggests that 164 words have the maximum value. Quite interestingly, the standard deviation of data is 9.99, which in turn implies that words are roughly larger than  $3.92-9.99$  and smaller than  $3.92+9.99$ . Now let us turn our attention to word length and its frequency:

**Table 3 Word length**

Value	Frequency	Proportion	Cumulative Proportion
2.0	14	0.01	0.01
3.0	92	0.063	0.073
4.0	201	0.138	0.211
5.0	195	0.134	0.344
6.0	218	0.15	0.494
7.0	179	0.123	0.617

8.0	167	0.115	0.731
9.0	132	0.091	0.822
10.0	101	0.069	0.891
11.0	63	0.043	0.934
12.0	27	0.019	0.953
13.0	18	0.012	0.965
14.0	21	0.014	0.979
15.0	6	0.004	0.984
16.0	6	0.004	0.988
17.0	2	0.001	0.989
18.0	2	0.001	0.99
19.0	3	0.002	0.992
20.0	2	0.001	0.994
21.0	2	0.001	0.995
25.0	2	0.001	0.997
26.0	1	0.001	0.997
28.0	1	0.001	0.998
31.0	1	0.001	0.999
35.0	1	0.001	0.999
38.0	1	0.001	1
Total	1458	1	

It is worth mentioning that the six-word expression has the highest frequency (218 tokens) and the highest proportion (0.15), as alluded to in Table 3. It should be noted, on the other hand, that the six-word expression is followed by the four-word expression. To be more specific, the frequency of

the four-word expression is 201 tokens (the second highest) and its proportion is 0.15. Notice that the four-word expression is followed by the five-word expression. More specifically, the frequency and proportion of the five-word expression are 195 tokens (the third highest) and 0.134. It is important to mention that the frequency of the seven-word expression is 179 tokens and that its proportion is 0.123. Now attention is paid to the descriptive properties of data analyzed by the software package NetMiner:

**Table 4 Descriptive properties**

Mean	Min	Max	Variance	Std. Dev.
7.089	2	38	11.283	3.359

Note that the mean (average) value of data is 7.089. This in turn suggests that the average 7.089-word expression occurred in 27 pieces of BBC news. It is worth pointing out, on the other hand, that the minimum value of words is 2 (the two-word expression) and that the maximum value of words is 38 (the thirty eight-word expression). Quite interestingly, the standard deviation of data is 3.359, which in turn means that words are roughly larger than  $7.089-3.359$  and smaller than  $7.089+3.359$ . Now let us turn our attention to the number of links. In Table 5, the so-called frequency refers to the number of links:

**Table 5 Number of links**

Value	Frequency	Proportion	Cumulative Proportion
1.0	1287	0.778	0.778
2.0	276	0.167	0.944
3.0	29	0.018	0.962
4.0	34	0.021	0.982
5.0	7	0.004	0.987

6.0	8	0.005	0.992
7.0	3	0.002	0.993
8.0	3	0.002	0.995
9.0	1	0.001	0.996
12.0	1	0.001	0.996
13.0	2	0.001	0.998
14.0	1	0.001	0.998
15.0	2	0.001	0.999
143.0	1	0.001	1
Total	1655	1	

In Table 5, the so-called value refers to the weight of links. When the weight of links is 1, the number of links is 1,287. That the weight of links is 1 indicates that word pairs are 1. As exemplified in Table 5, when the weight of links is 1, the number of links is the highest. On the other hand, when the weight of links is 2, the number of links is 276 (the second highest). When the weight of links is 3, there is a sharp decline in the number of links (29).

Now let us turn to Table 6. Table 6 shows information on degree. The term degree refers to the number of neighboring nodes directly linked. In-degree refers to the number of nodes directly linked to a node and the link(s) comes inside towards a node (nodes). On the other hand, the same can be said of out-degree, but the link(s) goes outside towards a node (nodes):

**Table 6 Distribution of degree**

Measures	Value	
	In-Degree	Out-Degree
Sum	1655	1655
Mean	1.352	1.352
Std. Dev.	2.43	2.426

<b>Min.</b>	0	0
<b>Max.</b>	45	43
<b># of Isolate</b>	0	
<b># of pendent</b>	853	
<b>Inclusiveness (%)</b>	100%	

It is important to note that the total number of in-degree and out-degree is 1,655, respectively. Quite interestingly, the mean (average) value of in-degree and out-degree analyzed by the software package NetMiner is 1.352, respectively. The standard deviation of in-degree is 2.43, which in turn means that in-degree is roughly larger than  $1.352-2.43$  and smaller than  $1.352+2.43$ . On the other hand, the standard deviation of out-degree is 2.426, which in turn indicates that out-degree is roughly larger than  $1.352-2.426$  and smaller than  $1.352+2.426$ . Interestingly, the minimum value of in-degree and out-degree is 0, which indicates that the number of links is 0. On the other hand, the maximum value

It is important to mention that as indicated in Table 7, the number of isolated nodes is 0. The term *transmitter* refers to the number of nodes sending a link (links). As indicated in Table 7, the total number of *transmitter* is 302. It should also be pointed out that the term *receiver* refers to the number of nodes receiving a link (links). As exemplified in Table 7, the total number of *receiver* is 324. The term *carrier* indicates that a node

of in-degree is 45, which implies that the maximum number of links is 45, whereas the maximum value of out-degree is 43, which indicates that the maximum number of links is 43. Perhaps it is worthwhile pointing out that the number of isolated links is 0. This in turn indicates that links look like a pendent (a necklace). It is worth noting that the number of the pendent is 853, as illustrated in Table 6. Finally, it is interesting to consider inclusiveness (the term of NetMiner) that refers to the proportion of linked nodes. As exemplified in Table 6, the proportion of linked nodes (inclusiveness) is 100%. Now attention is paid to the number and type of nodes:

**Table 7 Number and type of nodes**

Isolate	Transmitter	Receiver	Carrier	Ordinary
0	302	324	302	296

receives a link from another node and gives a link to it. Note that in this case, in-degree is 1 and out-degree is also 1. Quite interestingly, the total number of *carrier* is 302. It must be noted, on the other hand, that the number of the rest (*ordinary*) is 296. Now let us turn our attention to the network properties of data analyzed by the software package NetMiner:

**Table 8 Network properties**

	# of Links : O(m)	Density : O(m)	Average Degree : O(m)	# of Components(Weak) : O(m)	# of Components(Strong) : O(m)
<b>Word Network (Sentences)</b>	1,655	0.001	1.352	63	890

It is worth observing that the number of links occurred in 27 pieces of BBC news is 1,655 and that their density is 0.001. As indicated in Table 8, average degree is 1.352. The term *degree* refers to the number of neighboring nodes. It is particularly noteworthy that the number of *components (weak)*

is 63. The term *components (weak)* indicate that there is a link between two nodes. On the other hand, the term *components (strong)* indicate that there are links in two directions. That is to say, a node gives a link to another node and the node receives a link from another node. Quite

interestingly, the number of *components (strong)* is 890, which in turn suggests that the number of *components (strong)* is fourteen times higher that of

*components (weak)*. Finally, attention is also paid to the network properties of data analyzed by the software package NetMiner:

**Table 9 Network properties**

	Reciprocity(Arc) : O(m)	Reciprocity(Dyad) : O(m)	Transitivity : O(nm)	Clustering Coefficient : O(n^3)	Mean Distance : O(nm)
<b>Word Network (Sentences)</b>	0.028	0.014	0.02	0.053	6.272

Most importantly, the term *reciprocity (Arc)* refers to the proportion of the number of reciprocal links. On the other hand, the term *reciprocity (Dyad)* refers to the proportion of the number of reciprocal pairs (node pairs). Quite interestingly, *reciprocity (Arc)* is 0.028, whereas *reciprocity (Dyad)* is 0.014. When it comes to the term *transitivity*, things are different. It indicates that if a link goes from A to B and a link goes from B to C, then a link goes from A to C. The term *transitivity* takes place among three nodes. Simply put, my friend’s friend is my friend. As shown in Table 9, *transitivity* is 0.02. It is interesting to consider the term *clustering coefficient*. It refers to the probability value of associating a node with neighboring nodes. In other words, *clustering coefficient* is called *ego network density*. The term *ego* refers to a node located in a

central place and a node (nodes) linked with *ego* is called *alter*. Most importantly, the value of *clustering coefficient* is from 0 to 1. As illustrated in Table 9, the value of *clustering coefficient* is 0.053. Finally, the term *mean distance* is called *Average Path Length*. As shown in Table 9, the mean distance is 6.272. The term *path* refers to a trail without the overlap of nodes. It is like a navigator. The navigator never leads us to the passed place.

**2.2. Word cloud**

In what follows, we aim at providing a word cloud that represents 27 pieces of BBC news. Note that major keywords occur in different sizes. Note that each keyword is different in size, depending on the degree of importance:



**Figure 1 Word cloud**

It is interesting to observe that the noun *Game* is the biggest in size. This in turn implies that the noun *Game* is deemed to be the most significant and important. Note that the noun *Game* is followed by the noun *Squid*. More specifically, the noun *Squid* is the second biggest in size, which we can think of as the second most important. It is interesting to consider the noun *show*. When it comes to it, it is the third biggest. This amounts to saying that it is the third most significant. Finally, it is probably worthwhile pointing out that the noun *Netflix* ranks fourth. Simply put, it is the fourth biggest. We take this fact as meaning that it is the fourth most significant in 27 pieces of BBC news.

It seems thus reasonable to conclude that the noun *Game* is assumed to be the most significant, followed by the noun *Squid*, the noun *show*, and the noun *Netflix*, in that order.

### 2.3. Topic information

In the following, we aim to consider 12 topics that constitute 27 pieces of BBC news. Note that major keywords consist of 12 topics, which in turn constitute 27 pieces of BBC news:

**Table 10 Topic information**

	<b>1st Keyword</b>	<b>2nd Keyword</b>	<b>3rd Keyword</b>	<b>4th Keyword</b>	<b>5th Keyword</b>
<b>Topic-1</b>	Parasite	film	Oscars	picture	series
<b>Topic-2</b>	show	Squid	Credit	money	Game
<b>Topic-3</b>	show	world	Netflix	month	work
<b>Topic-4</b>	time	Parasite	Oscars	life	director
<b>Topic-5</b>	year	person	country	South Korea	film
<b>Topic-6</b>	Game	Squid	culture	pop	Netflix
<b>Topic-7</b>	film	language	subtitle	Parasite	audience
<b>Topic-8</b>	Netflix	content	service	Game	subtitle
<b>Topic-9</b>	person	viewer	life	series	K-dramas
<b>Topic-10</b>	world	viewer	translation	country	way
<b>Topic-11</b>	show	parent	series	school	teacher
<b>Topic-12</b>	game	series	child	person	group

It is worth mentioning that the keywords *Parasite*, *film*, *Oscars*, *picture*, and *series* consist of topic 1.

As can be seen from Table 10, the noun *Parasite* is the 1<sup>st</sup> keyword, thereby indicating that it is the most occurred one. It is important to mention, on the other hand, that the keywords *show*, *world*, *Netflix*, *month*, and *work* constitute topic 3. As exemplified in Table 10, the noun *show* is the 1<sup>st</sup> keyword, which we can think of as the most frequently used one. It is interesting to observe topic 8. Quite interestingly, the keywords *Netflix*, *content*, *service*, *Game*, and *subtitle* are made up of

topic 8. Interestingly, the noun *Netflix* occurred as the 1<sup>st</sup> keyword, which is considered to be the most widely used one in topic 8. Finally, it should also be noted that topic 9 contains the keywords *person*, *viewer*, *life*, *series*, and *K-dramas*. In this topic, the noun *person* is deemed to be the most frequently used one.

Now attention is paid to the frequency of each topic:

**Table 11 Topic frequency**

	# of sentences
<b>Topic-1</b>	46
<b>Topic-2</b>	32
<b>Topic-3</b>	41
<b>Topic-4</b>	36
<b>Topic-5</b>	38
<b>Topic-6</b>	97
<b>Topic-7</b>	72
<b>Topic-8</b>	39
<b>Topic-9</b>	53
<b>Topic-10</b>	29
<b>Topic-11</b>	34
<b>Topic-12</b>	89

It must be stressed that topic 6 occurred in 97 sentences (the highest). As shown in Table 10, the keywords *Game*, *Squid*, *culture*, *pop*, and *Netflix* are made up of topic 6. This in turn indicates that these keywords are the most widely used ones in 27 pieces of BBC news. It is interesting to consider topic 12. With respect to topic 12, it is interesting to note that it appeared in 89 sentences (the second highest). As alluded to in Table 12, the keywords

*game*, *series*, *child*, *person*, and *group* constitute topic 12. It must also be emphasized that topic 7 occurred in 72 sentences (the third highest). As can be seen from Table 10, the keywords *world*, *viewer*, *translation*, *country*, and *way* consist of topic 7. It seems thus reasonable to assume that these five keywords are considered to be the third most frequently used ones in 27 pieces of BBC news. Finally, it is interesting to observe topic 9. When it

comes to topic 9, it appeared in 53 sentences (the fourth highest), thus implying that the five keywords *person*, *viewer*, *life*, *series*, and *K-dramas* are the fourth most widely used ones. It can

thus be inferred that topic 6 is the most occurred one in 27 pieces of BBC news, followed by topic 12, topic 7, and topic 9, in descending order.

#### 2.4. Major words

This section is focused on looking into the frequency of major nouns occurred in 27 pieces of

BBC news. Our list was cut off in the top 30 for the reason of space:

**Table 12 Major words**

Number	Word	Frequency
1	Game	137
2	Squid	127
3	show	104
4	Netflix	90
5	game	73
6	person	69
7	series	64
8	film	61
9	Parasite	43
10	child	41
11	world	35
12	language	35
13	audience	34
14	year	33
15	subtitle	33
16	success	31
17	TV	26
18	content	25
19	viewer	24
20	story	24
21	country	22
22	way	19
23	parent	19
24	life	19
25	drama	19
26	platform	19





thereby having lower closeness centrality than the noun *Squid*. It is interesting to observe the noun *show* in the map. Also, it is near to an edge, as compared with the noun *Squid*, thus indicating that it counts as less significant. For the map of

### 3. Conclusion

To sum up, we, by using the software package NetMiner, have analyzed 27 pieces of BBC news regarding the movie *Squid Game*. In section 2.1, we have provided information on the structural and descriptive properties of data analyzed by the software package NetMiner. More specifically, we have provided the number of isolated links and the number of a pendent. Also, we have provided network properties such as the number of links, their density, average degree, and the number of components. We have also provided network properties such as reciprocity (Arc and Dyad), transitivity, clustering coefficient, and a mean distance of links. In section 2.2, we have shown that in the so-called word cloud, the noun *Game* is

synonyms and keywords, see Kang (2022a, 2022b, 2022d, 2022d, 2023a, 2023b). We thus conclude that the noun *Squid* has the highest closeness centrality.

deemed to be the most significant, followed by the noun *Squid*, the noun *show*, and the noun *Netflix*, in that order. In section 2.3, we have argued that topic 6 is the most occurred one in 27 pieces of BBC news, followed by topic 12, topic 7, and topic 9, in descending order. In section 2.4, we have maintained that the noun *Game* is the most widely used one in 27 pieces of BBC news, followed by the noun *Squid*, the noun *show*, and the noun *Netflix*, in that order. In section 2.5, we have further argued that the noun *Squid* has the highest degree centrality, followed by the noun *show*, and the noun *success*. In section 2.6, we have shown that the noun *Squid* has the highest closeness centrality.

### References

- [1] Kang, N. (2022a). A Comparative Analysis of Search for and Look for in Four Corpora. *Advances in Social Sciences Research Journal*, 9(3), 168-178.
- [2] Kang, N. (2022b). A Comparative Analysis of Impressed by and Impressed with in Two Corpora. *Theory and Practice in Language Studies*, 12(5), 819-827.
- [3] Kang, N. (2022c). On Speak to and Talk to: A Corpora-based Analysis. *Theory and Practice in Language Studies*, 12(7), 1262-1270.
- [4] Kang, N. (2022d). On Speak with and Talk with: A Corpora-based Analysis. *International Journal of Social Science and Human Research*, 5(8), 3354-3360.
- [5] Kang, N. (2023a). K-Pop in BBC News: A Big Data Analysis. *Advances in Social Sciences Research Journal*, 10(2), 156-169.
- [6] Kang, N. (2023b). K-Dramas in Google: A NetMiner Analysis. *Transaction on Engineering and Computing Sciences*, 11(1), 193-216.