We declare that this material, which we now submit for assessment, is entirely our own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of our work. We understand that plagiarism, collusion, and copying are grave and serious offences in the university and accept the penalties that would be imposed should I engage in plagiarism, collusion or copying. This assignment, or any part of it, has not been previously submitted by me/us or any other person for assessment on this or any other course of study.

Link Analysis

Amirreza Dashti Genave

amirreza.dashtigenave@studenti.unimi.it, Matriculation number: 963404

Mohammadhossein Jafari

mohammadhossein.jafari@studenti.unimi.it, Matriculation number: 964548

September 2022

Dataset Description

IMDb is a famous name in the movie Industry. On its website, it is possible to retrieve a wide range of information about movies. In this analysis, we use the IMDb dataset, which has been realized based on IMDb's non-commercial licensing. The Dataset has been published on the Kaggle website. The dataset is in the category of large datasets. It contains sub-datasets which are title.akas.tsv.gz, title.basics.tsv.gz, title.principals.tsv.gz, and title.ratings.tsv.gz. The raw data has been stored in the format of tab-separated value (TSV) files, which is a text format. In TSV files, data is stored in a table structure. Each record of the table is stored as one line in the text file. The purpose of the task is to create a ranking system between movies. For doing this activity it is needed to create a graph whose nodes are movies, and nodes are connected by edges if between two nodes there is at least one common actor or actress. It is possible to create the graph by using the data which is stored in two sub_datasets which are title.basics.tsv, and title.principals.tsv. The dataset title.basics.tsv is read and its separator is a tab. It has the following columns:

- Tconst which is an alphanumeric unique identifier of the title
- TitleType which is the type of the title such as movie, short, tvseries, tvepisode, video and etc.
- PrimaryTitle which is the more popular title / the title used by the filmmakers on promotional materials at the point of release
- OriginalTitle – original title in the original language
- StartYear (YYYY) – represents the release year of a title
- RuntimeMinutes – primary runtime of the title, in minutes
- Genres (string array) – includes up to three genres associated with the title

The title.basics dataset has 9202163 rows. The data is read during the execution of the code from the Kaggle API. The data is also updated weekly.

More than 75.4 percent of the data is composed of TV Episodes. The second rank is for short movies with 9.6 percent, and the third rank is for movies with 6.7 percent. The other types such as videos, tv movies are about 8 percent of the whole dataset.

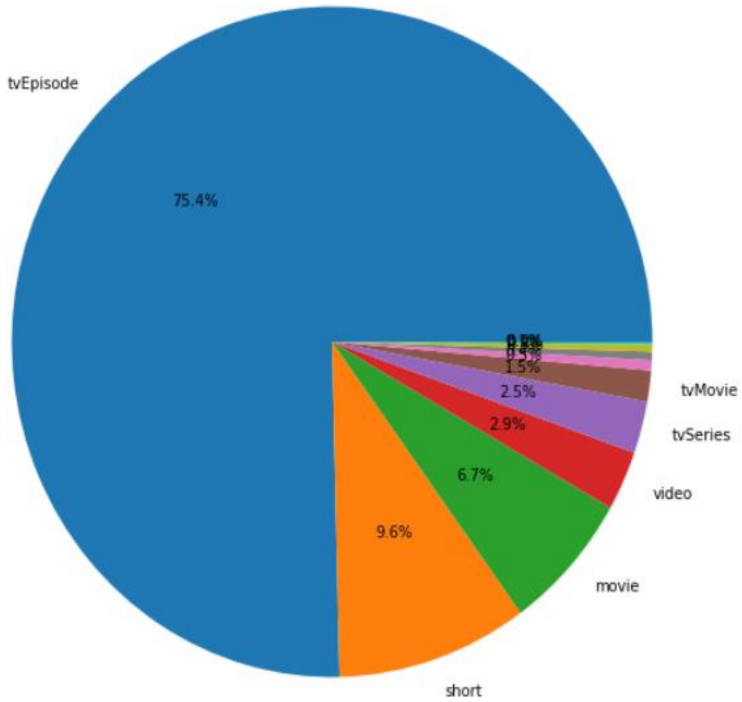

Figure 1

The dataset title.principals.tsv has information about principal cast/crew. It is read and its separator It has the following columns:

- Tconst which is an alphanumeric unique identifier of the title
- Ordering (integer) which is a number to uniquely identify rows for a given titleId
- Nconst (string) which is an alphanumeric unique identifier of the name/person.
- Category(string) which is the category of job that person was in.
- Job (string) includes the specific job title if applicable, else.
- Characters (string) - the name of the character played if applicable

Dataset Preprocessing

In order to implement the algorithm, firstly, it is needed to make the nodes of the graph. The first map function is implemented, and each record of the file title_basics is reduced to just two columns which are tconst and titleType. The created variable is named nodes, and the output is filtered by selecting titleType= movie. The total number of rows that are related to movies is 620463. In the second map function, the RDD node is limited to just the Tconst columns.

The Considered Algorithms and Their Implementations

The PageRank algorithm is a general concept, and there are different approaches to implementing it, and these different approaches can produce different results.

One large category of PageRank algorithms is based on probability. In other words, the result of the probability-based algorithm is a probability distribution, and a probability is assigned to each node.

There is also another variant of PageRank Algorithm generating float numbers that can be higher than 1 such as 38.28. It is important to mention that the first PageRank Algorithm which was proposed by Larry Page and Sergey Brin was based on this approach. Therefore, it is called the Historic Approach in this paper.

It is important to consider that two parameters can influence the finishing of the PageRank Algorithms. These two parameters are the **number of iterations** and the **number of Tolerances**.

In this paper, three different PageRank Algorithms are implemented, and their results are compared.

Historic Approach

Firstly, the initial PageRank Algorithm which is named the Historic Approach is implemented. The below formula demonstrates how the Historic PageRank is calculated

$$PR(A) = 1 - d + d\left(\frac{PR(B)}{L(B)} + \frac{PR(C)}{L(C)} + \frac{PR(D)}{L(D)} + \cdots\right)$$

Formula 1

PR(A) is the PageRank score relating to node A, and L(B) is the total number of outbound edges from node B. PR(B) is the PageRank score relating to node B.

d is the damping factor. The damping factor shows with what probability the surfer continues its random surfing. The teleportation probability is (1- damping factor). The teleportation probability says what the probability of transferring a surfer is from one node to another node randomly. **This damping factor is used to reduce the negative consequences of Dead Ends and Spider Traps**. Different research has been done to choose an appropriate number for the damping factor and they suggest 0.85 as the number for it.

In order to implement the Historic Approach, the GraphFrames package is used. GraphFrames is a package that has been developed for DataFrame-based Graphs, and it is used in Apache Spark. In the Historic Approach, the sum of the PageRank score of all nodes is equal to the total number of nodes. In the experimental part of the task, it is shown that the sum of PageRank scores is equal to the sum of all nodes.

Probabilistic Approach

Secondly, the Probabilistic PageRank Algorithm is implemented. In this approach, the sum of the probability of all nodes is equal to 1. The biggest advantage of this category is that it gives a decent criterion saying the more the probability is near 1, the more important the node is. This is the formula of the Probabilistic Approach.

$$PR(A) = \frac{1-d}{N} + d\left(\frac{PR(B)}{L(B)} + \frac{PR(C)}{L(C)} + \frac{PR(D)}{L(D)} + \cdots\right)$$

Formula 2

In this approach, the PageRank score of each node is divided by the total number of nodes, and then it is assigned to each node. The probabilistic approach is implemented by the networkX package. NetworkX is a Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks. It will be evaluated in the experimental part whether the sum of PageRank scores of all nodes is equal to 1.

Third Approach

A third PageRank Algorithm is implemented from scratch, and it is named the Third Algorithm. The third method is presented to rank the nodes of the mentioned graph and implement the PageRank from scratch in a distributed way. The method compute_page_rank is responsible to compute the PageRank. It takes the arguments below:

1. edges_rdd: an RDD containing all the edges of the graph
2. nodes_count: a number representing the number of vertices
3. damping_factor: a probability to avoid dead-ends and the spider trap
4. max_iter: an integer to determine the maximum number of matrix multiplications
5. Tolerance: a measurement unit to calculate the euclidian distance between each step

In the beginning, the method calculates the out-degree of each node by the countByKey method. Then, we calculate 1/out-degree for each element of the first argument and call it M. Then, we multiply the damping factor probability by the previous values and call it M_hat. Afterward, we build M transpose from M_hat to keep the source vertices in our calculation. The Matrix is ready by far.

Then, the vector is initialized by dividing one over the number of nodes. Before describing the core of the method, it is worth mentioning that in each step of the calculation we need to calculate the distance between the two last results which is handled by Euclidian distance.

Finally, we are ready to compute the matrices multiplication using the map-reduce job. For each out-degree value of matrix MT, they are multiplied by the corresponding row of the vector, and in the last step, the summation is applied. Then we assign the old value to the old_vector and the new value to the vector. We repeat this process until we reach the end of the maximum number of iterations, or the distance becomes lower than the Tolerance. The final results are stored in the vector variable.

After computing the PageRank algorithm with the third method, we are ready to compare our results with the first algorithm. To do so, we need to implement a method to calculate the similarity of the results of both algorithms. That is the purpose of the compute_similarity() method. Its arguments are as follows:

1. and 2. Two lists containing strings (vertices names) to be compared.

3. threshold: a number that configures the sensitivity of the method.

The compute_similarity method checks whether two identical nodes have the same place in both lists up to a certain threshold. For instance, with a threshold equal to two, it is acceptable to consider two identical nodes in 5^{th} place and 7^{th} place similar. The final result of similarity is the summation of the counter divided by the size of the lists. It returns a percentage. For instance, 0.8 means 80% of similarity between the results of the two algorithms.

One of the most important purposes of this task is to compare the results of the algorithms. The three algorithms are compared to check how the results of the algorithms are similar to each other.

The comparison is done in two different methods. In the first method, the result of algorithms for different numbers of iterations and Tolerance is calculated, and it is counted how many common movies exist in the top 20 movies of pair of Algorithms.

In the second method, the position of the movies in the top-20 results is important, and a threshold is defined for the position of the movie in the list. For example, if the threshold is 4, and a movie has the 1^{st} rank in the Historic approach and it has the 10^{th} rank in the Probabilistic Approach, it is not considered as similar while if instead of 10^{th} rank it was 3^{rd} rank, it was considered as similar.

Experiment Description

For doing the PageRank Algorithm the graph has to be made. An inner join is made between the title_principals and title_basics based on the tconst, which is the common column between the two datasets, with the aim of choosing only movies that are in both files. The result will be in a variable named sql_df.

In the next step, a new RDD is created named cast_id_to_movie_list. In this variable the nconst which is the unique identifier of each actor or actress is the key, and its value is a list of movies in which the person has played.

For creating this RDD, it is important to reduce the data to just people who are actors or actresses. Firstly, a map function is run on the sql_df to choose just the columns nconst, tconst, and category. Then, the method filter is used to choose just the actor or actress. With the help of a groupByKey and a map function, the cast_id_to_movie_list with the mentioned characteristics is made. It is the basis for making the nodes.

In the next step, a map function is executed to find the movies that have at least one common actor or actress. The new RDD is named movie_id_to_movie_id. Each element of movie_id_to_movie_id is a tuple of movies that share an actor or actress together. Therefore, if the person has played in 8 movies, there are 56 tuples in the related list of that special person. The method flatMap is run on the movie_id_to_movie_id. Therefore, its elements are just the tuples of movies that have at least one common actor or actress.

At this step, the initial graph has to be created. With the class “pyspark.sql.types.Row”, it is possible to create a row object. With the help of Row and doing a map function the vertices of the graph are made from the *nodes* variable with the command of v = nodes.map(row).toDF(). The toDF method provides a very concise way to create a Dataframe. The name of the variable containing all the vertices is named v.

The edges are also made by implementing the methods movie_id_to_movie_id.toDF([“src”, “dst”]) and the method dropDuplicates method. The name of the variable containing all the edges is named *e*. The DataFrame which is related to edges should contain two special columns: “src” (source vertex ID of edge) and “dst” (destination vertex ID of edge).

By having both the vertices and edges, it is tried to make the needed graph named g. The type of the variables v and e is pyspark.sql.dataframe. GraphFrame is a package which has been developed with Apache Spark. It can be used to provide DataFrame-Based Graphs. By considering that the edges and the vertices are ready, it is possible to create the graph with the help of GraphFrames. The created graph is named g.

The GraphFrames Library has a method which is named PageRank. Then we rank the verticies based on their importance. By doing that the PageRank which is based on the Historic Approach is made.

The PageRank which is based on the Probabilistic Approach is implemented by the library NetworkX. The PageRank algorithm running by NetworkX receives the graph *g* as its input. The graphs in NetworkX are directed. If the graph is undirected, it will be changed to a directed Graph. This matter is done by converting each edge to two edges.

- The damping factor in this algorithm is alpha and it has a value in the range from 0 to 1. It is possible to choose the maximum number of iterations of the power method in the eigenvalue solver, it is an optional value and it has to be an integer. This parameter is named max_iter.
- The convergence in the power method can be monitored by an optional parameter named tol. It is the error Tolerance that analyzes the convergence.
- It is possible to assign a value as a PageRank to each node at the start of the algorithm. The assignment is done by a dictionary. The name of this parameter is named nstart.
- The algorithm returns a dictionary that has the node names and the PageRank Value.
- The algorithm finishes when the maximum number of iterations is finished or the error Tolerance of graph_node_number * tol is reached.
- If the algorithm is not able to reach convergence based on the specified value for the Tolerance in the specified value for the maximum number of iterations, the algorithm returns **PowerIterationFailedConvergence.**
- The networkX has PageRank_numpy, PageRank_scipy, networkx.PageRank Algorithms. Each of the three Algorithms has a different approach to dealing with the task.

Networkx.PageRank_scipy() solves the power method by implementing the SciPy sparse-matrix. It has two parameters which can be tuned and can control the accuracy, which are tol and maix_iter

At the same time, for the aim of calculating the largest eigenvalue/eigenvector of the Google matrix, the networkx.PageRank() uses a pure python implementation of the power method. The networkx.PageRank() has also the two parameters tol and maix_iter.

On the other hand, the networkx.PageRank_numpy() is a a Numpy (full) Matrix implementation to compute the largest eigenvalue and eigenvector.

Comments and Discussion On The Experimental Results

The three algorithms are implemented correctly, and they produce their own results. It is shown that the sum of the PageRank Score of all edges is equal to the total number of edges, and the sum of the PageRank score of all nodes in the Probabilistic Approach is equal to 1. As it has been explained the compute_similarity method at the end of the code checks whether two identical nodes have the same place in both lists up to a certain threshold. For instance, with a threshold equal to two, it is acceptable to consider two identical nodes in 5^{th} place and 7^{th} place similar. The final result of similarity is the summation of the counter divided by the size of the lists. It returns a percentage.

For instance, 0.65 means 65% of similarity between the results of the two algorithms. Here the similarity is 80 percent.

On the other hand, Table 1 and Table 2 compare the number of common movies among the top-20 movies of the different PageRank Algorithms by considering different numbers of Tolerance and the number of iterations. For example, at the Tolerance level of 0.01, there are 9 movies that are common between the top-20 movies of Probabilistic, and Historical Approaches. The number for common movies is at the Tolerance level of 0.001 and it increases to 18 at the Tolerance level.000001.

As it can be seen when the criteria of finishing the PageRank Algorithm is the maximum number of iterations, there is higher number of common movies between the Historic and Probabilistic Mehtod.

	Iteration = 10	Iteration = 20
Historic and Probabilistic	19	19
Probabilistic and Third	18	16
Historic and Third	19	18

Table 1

	Tolerance = 0.01	Tolerance = 0.001	Tolerance = 0.000001
Probabilistic and Historic	9	9	18

Table 2

Table 3 shows the results of the Probabilistic and Historic Approach at the Tolerance level of 0.000001.

<table>
<tr><th>Historic Approach
0.000001 TOL</th><th>Probabilistic Approach</th></tr>
<tr><td>id pagerank</td><td>id pagerank</td></tr>
<tr><td>|tt14847122|29.953381455028303|</td><td>('tt14847122', 0.00024469300658203064),</td></tr>
<tr><td>|tt20566026|28.391604723407408|</td><td>('tt20566026', 0.00023126339790712348),</td></tr>
<tr><td>| tt6333648|25.727589260165345|</td><td>('tt6333648', 0.00021229002277163185),</td></tr>
<tr><td>|tt15743196|20.442618581192228|</td><td>('tt9788846', 0.0001666804271596129),</td></tr>
<tr><td>| tt9788846|20.204404322439228|</td><td>('tt15743196', 0.00016594513321545846),</td></tr>
<tr><td>|tt10026090|19.936663266769674|</td><td>('tt20197090', 0.00016395132078084518),</td></tr>
<tr><td>|tt20197090|19.785578123983868|</td><td>('tt10026090', 0.00016269786551075468),</td></tr>
<tr><td>|tt21038574| 19.42100828513236|</td><td>('tt21038574', 0.00015916013965840814),</td></tr>
<tr><td>|tt15356682| 19.35684727518884|</td><td>('tt15356682', 0.00015837113541489342),</td></tr>
<tr><td>|tt12729186|19.251516110861743|</td><td>('tt16262852', 0.0001580670617079871),</td></tr>
<tr><td>|tt16262852|19.153356767589887|</td><td>('tt15500386', 0.00015760158731221),</td></tr>
<tr><td>|tt15500386| 19.13871902121548|</td><td>('tt12729186', 0.00015740172699506418),</td></tr>
<tr><td>|tt18258932|19.136216796409425|</td><td>('tt18258932', 0.0001570336899847899),</td></tr>
<tr><td>|tt21608688| 18.84580754384966|</td><td>('tt13849554', 0.00015507363761541455),</td></tr>
<tr><td>|tt19864468|18.839234557834942|</td><td>('tt14536348', 0.00015213731621393473),</td></tr>
<tr><td>|tt21228982|18.763000314729318|</td><td>('tt19864468', 0.0001506186511817291),</td></tr>
<tr><td>|tt13849554|18.632781745559665|</td><td>('tt21608688', 0.0001493138709634496),</td></tr>
<tr><td>|tt14536348|18.411016172307324|</td><td>('tt21352732', 0.00014871030256146859),</td></tr>
<tr><td>|tt16710748| 18.06107350085137|</td><td>('tt21228982', 0.00014831191390742317),</td></tr>
<tr><td>|tt18951894|18.026140923781014|</td><td>('tt21056872', 0.00014775974533804483),</td></tr>
</table>

Table 3

A small sample of the original graph with a cool-warm color map that shows the value the of pagerank has been generated. In the generated image below, nodes(movies) with more connecting edges are warmer.

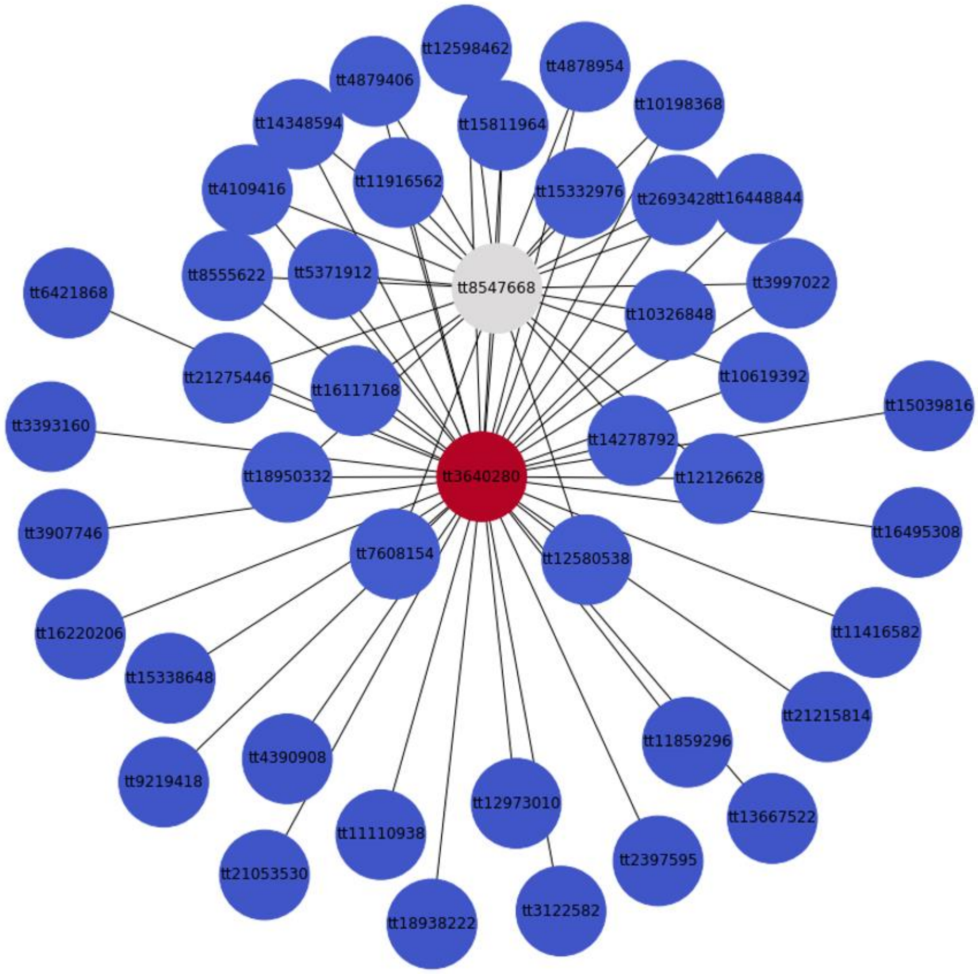

Figure 2

A Collection of tables of results are demonstrated at the end of the report.

How The Proposed Solution Scales Up With Data Size

The entire process of making the data ready to feed the PageRank algorithm has been done in resilient distributed datasets (RDD). In some sections of the code, the RDD is converted to SQL Dataframe to make the process more convenient. According to the Spark documentation, the computational costs of both approaches are identical. Moreover, the probability matrix and the power method which is used to calculate the PageRank are implemented by map-reduce jobs.

The only part of the calculation which is not considered as an RDD is the PageRank vector. In a worst-case scenario, the size of this vector could roughly reach 8 * e9 bytes which is still much less than available RAM in nowadays clusters.

The solution scales up according to dataset size. In order to guarantee that the solution can work properly for larger datasets. The type “movie” is replaced with the type “short”. The volume of the “***short***” type is 43 percent higher than the volume of the ***“movie”*** type. The proposed Algorithms are implemented on the “short” type, and it is shown that the Algorithms work well also for larger Datasets. By considering Historic and the Probabilistic PageRank methods, the number of common short movies is 19 out of 20, which is an accuracy of 95 percent. The tables below show the ranking Table for the Historic and the ProbabilisticAlgorithms. These results are for the maximum number of iterations of 10.

Historic Approach		Probabilistic Approach
id	pagerank	
tt4745692	38.129513875038946	[('tt4745692', 0.0002163202093671938),
tt11916562	37.902554176792556	('tt3886736', 0.0002119182217079988),
tt3886736	36.644535032104656	('tt11916562', 0.0002077988514825338),
tt4968564	34.427414951191594	('tt5918332', 0.00018751721057302289),
tt10765798	33.583327860112036	('tt10765798', 0.00018525163612327557),
tt5918332	33.511759683987755	('tt4968564', 0.00018430113577553323),
tt13114794	32.32537378554175	('tt13114794', 0.00017800891455590747),
tt21215814	29.896344240368087	('tt21215814', 0.00016932777192667694),
tt14132588	29.844176814119663	('tt12399882', 0.00016801499272454175),
tt9512094	29.64352220668977	('tt11776266', 0.00016505671892582982),
tt12399882	29.564646712909955	('tt9512094', 0.00016077731867347449),
tt13950392	29.508764495925448	('tt6496598', 0.00016039697063655232),
tt14416808	29.245259849291866	('tt13950392', 0.00015984858831305464),
tt12126628	28.822122299149537	('tt14132588', 0.0001592065264466762),
tt11776266	28.600410836287512	('tt12126628', 0.00015743126715840616),
tt10921118	28.58290232053236	('tt10921118', 0.00015666348459495585),
tt12709774	28.34047313295296	('tt14416808', 0.00015602474027694347),
tt11075800	28.16429226028806	('tt11075800', 0.0001542162388808778),
tt3643390	28.154874648541494	('tt3643390', 0.0001540468186092864),
tt11416582	28.001821591730373	('tt12709774', 0.00015383212108394708)]

Result Tables

Historic Approach 0.01 TOL		Probabilistic Approach	
id	pagerank	id	pagerank
tt14847122	26.813650608093717	('tt6333648',	0.00028352451425149393),
tt20566026	25.521786621151083	('tt13836906',	0.00021833049692877835),
tt6333648	22.726614627288622	('tt20566026',	0.0002097135595176557),
tt9788846	18.574749271934742	('tt21056872',	0.0001940728444487843),
tt20197090	18.088056232967507	('tt8005490',	0.00018600342848463147),
tt15743196	18.018763293792702	('tt8932070',	0.00018542704163004912),
tt10026090	17.679964512432406	('tt14847122',	0.0001842009391867033),
tt15356682	17.64806767880321	('tt14614958',	0.00018020812065583082),
tt21038574	17.547161894289548	('tt20197090',	0.0001774729030368128),
tt16262852	17.50437056348854	('tt10026090',	0.00017702935112123667),
tt15500386	17.379482488761553	('tt12729186',	0.00017559552871045887),
tt12729186	17.281141310278517	('tt13849554',	0.00017521676020601905),
tt18258932	17.14914271203474	('tt13373822',	0.00017470624613481753),
tt19864468	17.027866271626493	('tt20770946',	0.00017332291768381992),
tt21608688	16.99680093922553	('tt21255488',	0.00017332291768381992),

|tt13849554|16.834017750970208| |tt14536348|16.724522931793654| |tt21352732|16.467848178870284| |tt21228982| 16.38011017433762| |tt16710748|16.342219936078866|	('tt21228982', 0.00017258830505091576), ('tt14626330', 0.0001712479250073235), ('tt9788846', 0.00017027779856116932), ('tt8795670', 0.00016986459655632586), ('tt12392116', 0.0001636041485665546)

Historic Approach 0.001 TOL	Probabilistic Approach
id pagerank	id pagerank
|tt14847122|29.638563593028504| |tt20566026| 28.10064293811751| | tt6333648|25.424870993992972| |tt15743196| 20.21689507174325| | tt9788846|20.026307832253092| |tt10026090|19.710805339195183| |tt20197090|19.613394143708735| |tt21038574| 19.23380687495996| |tt15356682|19.180700284932175| |tt12729186|19.064770995744244| |tt16262852| 18.9836158856361| |tt15500386|18.961247413835835| |tt18258932| 18.95013725560532| |tt19864468|18.663994388340257| |tt21608688| 18.66233592025297| |tt21228982|18.540263719406834| |tt13849554|18.452913969001333| |tt14536348|18.249253770637736| |tt16710748|17.893486980176753| |tt21352732|17.841540060502087|	'tt6333648', 0.00028352451425149393), ('tt13836906', 0.00021833049692877835), ('tt20566026', 0.0002097135595176557), ('tt21056872', 0.0001940728444487843), ('tt8005490', 0.00018600342848463147), ('tt8932070', 0.00018542704163004912), ('tt14847122', 0.0001842009391867033), ('tt14614958', 0.00018020812065583082), ('tt20197090', 0.0001774729030368128), ('tt10026090', 0.00017702935112123667), ('tt12729186', 0.00017559552871045887), ('tt13849554', 0.00017521676020601905), ('tt13373822', 0.00017470624613481753), ('tt20770946', 0.00017332291768381992), ('tt21255488', 0.00017332291768381992), ('tt21228982', 0.00017258830505091576), ('tt14626330', 0.0001712479250073235), ('tt9788846', 0.00017027779856116932), ('tt8795670', 0.00016986459655632586), ('tt12392116', 0.0001636041485665546),

Historic Approach 0.000001 TOL	Probabilistic Approach
id pagerank	id pagerank
|tt14847122|29.953381455028303| |tt20566026|28.391604723407408| | tt6333648|25.727589260165345| |tt15743196|20.442618581192228| | tt9788846|20.204404322439228| |tt10026090|19.936663266769674| |tt20197090|19.785578123983868| |tt21038574| 19.42100828513236| |tt15356682| 19.35684727518884| |tt12729186|19.251516110861743| |tt16262852|19.153356767589887| |tt15500386| 19.13871902121548| |tt18258932|19.136216796409425| |tt21608688| 18.84580754384966| |tt19864468|18.839234557834942| |tt21228982|18.763000314729318| |tt13849554|18.632781745559665| |tt14536348|18.411016172307324| |tt16710748| 18.06107350085137| |tt18951894|18.026140923781014|	('tt14847122', 0.00024469300658203064), ('tt20566026', 0.00023126339790712348), ('tt6333648', 0.00021229002277163185), ('tt9788846', 0.0001666804271596129), ('tt15743196', 0.00016594513321545846), ('tt20197090', 0.00016395132078084518), ('tt10026090', 0.00016269786551075468), ('tt21038574', 0.00015916013965840814), ('tt15356682', 0.00015837113541489342), ('tt16262852', 0.0001580670617079871), ('tt15500386', 0.00015760158731221), ('tt12729186', 0.00015740172699506418), ('tt18258932', 0.0001570336899847899), ('tt13849554', 0.00015507363761541455), ('tt14536348', 0.00015213731621393473), ('tt19864468', 0.0001506186511817291), ('tt21608688', 0.0001493138709634496), ('tt21352732', 0.00014871030256146859), ('tt21228982', 0.00014831191390742317), ('tt21056872', 0.00014775974533804483),

id |pagerank| Historic, iter = 10	id| pagerank| Probabilistic , iter = 10
| tt4745692| 38.12951387503893| |tt11916562| 37.90255417679254| | tt3886736| 36.64453503210465| | tt4968564| 34.42741495119159| |tt10765798| 33.58332786011202| | tt5918332| 33.51175968398774| |tt13114794|32.325373785541736| |tt21215814|29.896344240368077| |tt14132588|29.844176814119653| | tt9512094| 29.64352220668976| |tt12399882|29.564646712909944| |tt13950392|29.508764495925437| |tt14416808|29.245259849291855| |tt12126628|28.822122299149527| |tt11776266| 28.6004108362875| |tt10921118|28.582902320532348| |tt12709774|28.340473132952948| |tt11075800| 28.16429226028805| | tt3643390|28.154874648541483| |tt11416582|28.001821591730362|	('tt4745692', 0.0002163202093671938), ('tt3886736', 0.0002119182217079988), ('tt11916562', 0.0002077988514825338), ('tt5918332', 0.00018751721057302289), ('tt10765798', 0.00018525163612327557), ('tt4968564', 0.0001843011357755323), ('tt13114794', 0.00017800891455590747), ('tt21215814', 0.00016932777192667694), ('tt12399882', 0.00016801499272454175), ('tt11776266', 0.00016505671892582982), ('tt9512094', 0.0001607731867347449), ('tt6496598', 0.00016039697063655232), ('tt13950392', 0.00015984858831305464), ('tt14132588', 0.0001592065264466762), ('tt12126628', 0.00015743126715840616), ('tt10921118', 0.00015666348459495585), ('tt14416808', 0.00015602474027694347), ('tt11075800', 0.0001542162388808778), ('tt3643390', 0.0001540468186092864), ('tt12709774', 0.00015383212108394708]

id |pagerank| Historic, iter = 10	id| pagerank| Third , iter = 10
| tt4745692| 38.12951387503893| |tt11916562| 37.90255417679254| | tt3886736| 36.64453503210465| | tt4968564| 34.42741495119159| |tt10765798| 33.58332786011202| | tt5918332| 33.51175968398774| |tt13114794|32.325373785541736| |tt21215814|29.896344240368077| |tt14132588|29.844176814119653| | tt9512094| 29.64352220668976| |tt12399882|29.564646712909944| |tt13950392|29.508764495925437| |tt14416808|29.245259849291855| |tt12126628|28.822122299149527| |tt11776266| 28.6004108362875| |tt10921118|28.582902320532348| |tt12709774|28.340473132952948| |tt11075800| 28.16429226028805| | tt3643390|28.154874648541483| |tt11416582|28.001821591730362|	|[('tt11916562', 3.7873084898765935e-06), ('tt4968564', 3.584537078504016e-06), ('tt4745692', 3.5582565777718306e-06), ('tt10765798', 3.37191067188635e-06), ('tt3886736', 3.2665303501597442e-06), ('tt13114794', 3.2457170105959336e-06), ('tt5918332', 3.2151268161081943e-06), ('tt14416808', 3.1288112789106096e-06), ('tt14132588', 3.1100193171820667e-06), ('tt9512094', 3.0951663498970564e-06), ('tt10921118', 3.079555309656031e-06), ('tt11416582', 3.0692586906129624e-06), ('tt13950392', 3.040721069579688e-06), ('tt12126628', 2.9013227945781444e-06), ('tt12709774', 2.8978424835577967e-06), ('tt11075800', 2.891264983347567e-06), ('tt21215814', 2.8571663265419744e-06), ('tt12399882', 2.7974737395165135e-06), ('tt3643390', 2.742085462079488e-06), ('tt14069222', 2.71453076317109e-06)]

<table>
<tr><td>id |pagerank|
Historic, iter = 20</td><td>id| pagerank|
Probabilistic , iter = 20</td></tr>
<tr><td>|tt11916562| 39.60245884175969|
tt4745692	39.442991818003655
tt3886736	37.80073131259984
tt4968564	36.19073002151844
tt10765798	35.37464543176991
tt5918332	34.77857872337126
tt13114794	33.77550647509047
tt14132588	31.41995527539344
tt9512094	31.087462854691783
tt21215814	30.980978912141293
tt13950392	30.937774993113365
tt14416808	30.80134662924967
tt12399882	30.545687692929878
tt12126628	30.103351732781544
tt10921118	29.725833473910278
tt12709774	29.71092271080792
tt3643390	29.494070665436922
tt11075800	29.485127876320423
tt11776266	29.45431171448077
tt11416582|29.318910908517562|</td><td>[('tt4745692', 0.0002163202093671938), ('tt3886736', 0.0002119182217079988), ('tt11916562', 0.0002077988514825338), ('tt5918332', 0.00018751721057302289), ('tt10765798', 0.00018525163612327557), ('tt4968564', 0.00018430113577553323), ('tt13114794', 0.00017800891455590747), ('tt21215814', 0.00016932777192667694), ('tt12399882', 0.00016801499272454175), ('tt11776266', 0.00016505671892582982), ('tt9512094', 0.00016077731867347449), ('tt6496598', 0.00016039697063655232), ('tt13950392', 0.00015984858831305464), ('tt14132588', 0.0001592065264466762), ('tt12126628', 0.00015743126715840616), ('tt10921118', 0.00015666348459495585), ('tt14416808', 0.00015602474027694347), ('tt11075800', 0.0001542162388808778), ('tt3643390', 0.0001540468186092864), ('tt12709774', 0.00015383212108394708), ('tt13498286', 0.00015322735672170244),</td></tr>
</table>

<table>
<tr><td>id |pagerank|
Historic, iter = 20</td><td>id| pagerank|
Third , iter = 20</td></tr>
<tr><td>|tt11916562| 39.60245884175969|
tt4745692	39.442991818003655
tt3886736	37.80073131259984
tt4968564	36.19073002151844
tt10765798	35.37464543176991
tt5918332	34.77857872337126
tt13114794	33.77550647509047
tt14132588	31.41995527539344
tt9512094	31.087462854691783
tt21215814	30.980978912141293
tt13950392	30.937774993113365
tt14416808	30.80134662924967
tt12399882	30.545687692929878
tt12126628	30.103351732781544
tt10921118	29.725833473910278
tt12709774	29.71092271080792
tt3643390	29.494070665436922
tt11075800	29.485127876320423
tt11776266	29.45431171448077
tt11416582|29.318910908517562|</td><td>('tt11916562', 9.396179681208181e-07), ('tt4968564', 9.119836460585231e-07), ('tt10765798', 8.9009466162844e-07), ('tt4745692', 8.520172028341107e-07), ('tt13114794', 8.08202600046158e-07), ('tt14416808', 8.004274622201267e-07), ('tt14132588', 7.943039650824513e-07), ('tt9512094', 7.857680749692736e-07), ('tt5918332', 7.812505107772058e-07), ('tt3886736', 7.791872338657831e-07), ('tt11416582', 7.669803762898822e-07), ('tt13950392', 7.652730653514305e-07), ('tt10921118', 7.488906338049087e-07), ('tt11075800', 7.301035641969747e-07), ('tt12709774', 7.280448649377698e-07), ('tt12126628', 7.20684262036807e-07), ('tt14997306', 6.991207511122473e-07), ('tt15847776', 6.979525277802831e-07), ('tt21215814', 6.932526976475364e-07), ('tt3643390', 6.916039837891976e-07),</td></tr>
</table>

id |pagerank| Probabilistic, iter = 20	id| pagerank| Third , iter = 20
[('tt4745692', 0.0002163202093671938), ('tt3886736', 0.0002119182217079988), ('tt11916562', 0.0002077988514825338), ('tt5918332', 0.00018751721057302289), ('tt10765798', 0.00018525163612327557), ('tt4968564', 0.00018430113577553323), ('tt13114794', 0.00017800891455590747), ('tt21215814', 0.00016932777192667694), ('tt12399882', 0.00016801499272454175), ('tt11776266', 0.00016505671892582982), ('tt9512094', 0.00016077731867347449), ('tt6496598', 0.0001603969706365523 2), ('tt13950392', 0.00015984858831305464), ('tt14132588', 0.0001592065264466762), ('tt12126628', 0.00015743126715840616), ('tt10921118', 0.00015666348459495585), ('tt14416808', 0.00015602474027694347), ('tt11075800', 0.0001542162388808778), ('tt3643390', 0.0001540468186092864), ('tt12709774', 0.00015383212108394708), ('tt13498286', 0.00015322735672170244),	('tt11916562', 9.396179681208181e-07), ('tt4968564', 9.119836460585231e-07), ('tt10765798', 8.9009466162844e-07), ('tt4745692', 8.520172028341107e-07), ('tt13114794', 8.08202600046158e-07), ('tt14416808', 8.004274622201267e-07), ('tt14132588', 7.943039650824513e-07), ('tt9512094', 7.857680749692736e-07), ('tt5918332', 7.812505107772058e-07), ('tt3886736', 7.791872338657831e-07), ('tt11416582', 7.669803762898822e-07), ('tt13950392', 7.652730653514305e-07), ('tt10921118', 7.488906338049087e-07), ('tt11075800', 7.301035641969747e-07), ('tt12709774', 7.280448649377698e-07), ('tt12126628', 7.20684262036807e-07), ('tt14997306', 6.991207511122473e-07), ('tt15847776', 6.979525277802831e-07), ('tt21215814', 6.932526976475364e-07), ('tt3643390', 6.916039837891976e-07),

id| pagerank| Probabilistic , iter = 10	id| pagerank| Third , iter = 10
('tt4745692', 0.0002163202093671938), ('tt3886736', 0.0002119182217079988), ('tt11916562', 0.0002077988514825338), ('tt5918332', 0.00018751721057302289), ('tt10765798', 0.00018525163612327557), ('tt4968564', 0.00018430113577553323), ('tt13114794', 0.00017800891455590747), ('tt21215814', 0.00016932777192667694), ('tt12399882', 0.00016801499272454175), ('tt11776266', 0.00016505671892582982), ('tt9512094', 0.00016077731867347449), ('tt6496598', 0.0001603969706365523 2), ('tt13950392', 0.00015984858831305464), ('tt14132588', 0.0001592065264466762), ('tt12126628', 0.00015743126715840616), ('tt10921118', 0.00015666348459495585), ('tt14416808', 0.00015602474027694347), ('tt11075800', 0.0001542162388808778), ('tt3643390', 0.0001540468186092864), ('tt12709774', 0.00015383212108394708]	|[('tt11916562', 3.7873084898765935e-06), ('tt4968564', 3.584537078504016e-06), ('tt4745692', 3.5582565777718306e-06), ('tt10765798', 3.37191067188635e-06), ('tt3886736', 3.2665303501597442e-06), ('tt13114794', 3.2457170105959336e-06), ('tt5918332', 3.2151268161081943e-06), ('tt14416808', 3.1288112789106096e-06), ('tt14132588', 3.1100193171820667e-06), ('tt9512094', 3.0951663498970564e-06), ('tt10921118', 3.079555309656031e-06), ('tt11416582', 3.0692586906129624e-06), ('tt13950392', 3.040721069579688e-06), ('tt12126628', 2.9013227945781444e-06), ('tt12709774', 2.8978424835577967e-06), ('tt11075800', 2.891264983347567e-06), ('tt21215814', 2.8571663265419744e-06), ('tt12399882', 2.7974737395165135e-06), ('tt3643390', 2.742085462079488e-06), ('tt14069222', 2.71453076317109e-06)]

<table>
<tr><td>Historic Approach
20 ITER, 200 Movies</td><td>Probabilistic Approach</td></tr>
<tr><td>id pagerank</td><td>id pagerank</td></tr>
<tr><td>|tt11916562| 39.60245884175969|
tt4745692	39.442991818003655
tt3886736	37.80073131259984
tt4968564	36.19073002151844
tt10765798	35.37464543176991
tt5918332	34.77857872337126
tt13114794	33.77550647509047
tt14132588	31.41995527539344
tt9512094	31.087462854691783
tt21215814	30.980978912141293
tt13950392	30.937774993113365
tt14416808	30.80134662924967
tt12399882	30.545687692929878
tt12126628	30.103351732781544
tt10921118	29.725833473910278
tt12709774	29.71092271080792
tt3643390	29.494070665436922
tt11075800	29.485127876320423
tt11776266	29.45431171448077
tt11416582	29.318910908517562
tt19387356	28.83363581194299
tt10670022	28.681638793371878
tt14069222	28.116447466284278
tt14182380	27.86383142403571
tt15847776	27.22761202399111
tt17278774	27.12057222274021
tt14997306	27.113507849174532
tt13904218	27.097293529291054
tt12658830	27.029154194197076
tt6496598	26.75855479571765
tt12498996	26.66925963535965
tt17155092	26.613561565173868
tt2397595	26.385629033396754
tt12752880	26.32560331309736
tt21239560	26.02203480066064
tt13542474	25.85856515646123
tt13498286	25.791252392444637
tt15825262	25.75328979333861
tt8435252	25.635898159855746
tt11305742	25.450994699110897
tt12623450	25.390726631619273
tt15405486	25.247275691266598
tt5167166	25.18855998830985
tt1174971	25.146505894114252
tt16048054	25.11309483013625
tt12799846	25.09433517645111
tt9169200	24.893568160013853
tt4390908	24.823939731434944
tt16760028	24.808228593089893
tt15811964	24.802365836835165
tt12768728	24.71901262688493
tt2195857	24.65106988293704
tt11834062	24.626912346595375
tt11430948|24.511181520960026|</td><td>[('tt4745692', 0.0002163202093671938), ('tt3886736', 0.0002119182217079988), ('tt11916562', 0.0002077988514825338), ('tt5918332', 0.00018751721057302289), ('tt10765798', 0.00018525163612327557), ('tt4968564', 0.0001843011357755323), ('tt13114794', 0.00017800891455590747), ('tt21215814', 0.00016932777192667694), ('tt12399882', 0.00016801499272454175), ('tt11776266', 0.00016505671892582982), ('tt9512094', 0.00016077731867347449), ('tt6496598', 0.0001603969706365232), ('tt13950392', 0.00015984858831305464), ('tt14132588', 0.0001592065264466762), ('tt12126628', 0.00015743126715840616), ('tt10921118', 0.00015666348459495585), ('tt14416808', 0.00015602474027694347), ('tt11075800', 0.0001542162388808778), ('tt3643390', 0.0001540468186092864), ('tt12709774', 0.00015383212108394708), ('tt13498286', 0.00015322735672170244), ('tt13904218', 0.00015264265665299772), ('tt19387356', 0.00015242689782367857), ('tt11416582', 0.00015084335228583994), ('tt14182380', 0.0001499979232272674), ('tt10670022', 0.00014997397108241917), ('tt14069222', 0.0001476860273394416), ('tt2397595', 0.0001463559637863334), ('tt16450020', 0.00014632526022390662), ('tt5167166', 0.0001457807103193064), ('tt12498996', 0.0001452195510055958), ('tt17155092', 0.00014433345974199362), ('tt16048054', 0.00014328561892115702), ('tt1174971', 0.0001431154487860039), ('tt15847776', 0.0001417591495341676), ('tt17278774', 0.00014095415158886493), ('tt16760028', 0.00014062225218721013), ('tt4167532', 0.00013928131398403983), ('tt12658830', 0.0001386575390296042), ('tt4390908', 0.00013850764210050512), ('tt8435252', 0.00013838878990371257), ('tt12752880', 0.00013818419964414805), ('tt13542474', 0.00013809862990901342), ('tt12623450', 0.0001378899864140864), ('tt4581992', 0.000137776931016182), ('tt8586578', 0.000137614308845436), ('tt3143458', 0.00013640961950381222), ('tt13993652', 0.00013564855108873778), ('tt21239560', 0.00013557282449204187), ('tt14997306', 0.00013543734876016316), ('tt15825262', 0.00013496093112624262), ('tt4581420', 0.00013467450739077476), ('tt12768728', 0.0001344972332361274), ('tt9169200', 0.00013424480764509723),</td></tr>
</table>

<table>
<tr>
<td>| tt4167532|24.345271484651207|
| tt8586578|24.318209543919252|
|tt12439942| 24.275716761322|
tt11859296	24.201742377383813
tt4581992	24.00163248222558
tt2759098	23.940629670747438
tt16450020	23.894127765723233
tt15357686	23.82859632126881
tt3831118	23.798984343805714
tt7462288	23.73601060840957
tt10833362	23.578935329559375
tt14110888	23.563986945281837
tt18344978	23.475950693883235
tt15791002	23.342743069692492
tt14348594	23.278978747936417
tt15536618	23.211521731439838
tt2173216	23.168657248651286
tt5371912	23.118330465654733
tt21572042	23.101302133606797
tt10928904	22.979324361437907
tt15249688	22.95723290722756
tt11663228	22.88556536760316
tt3143458	22.871672614293065
tt20918100	22.867031988128648
tt6421868	22.849902258012524
tt12221126	22.835680837521704
tt13993652	22.742166291892698
tt11465042	22.731711795665063
tt12234316	22.715206799633574
tt11110938	22.697475347813658
tt2365971	22.674755666156006
tt14213908	22.66076434844518
tt7353622	22.47783029293113
tt12338246	22.408920266227483
tt21141138	22.350678243679848
tt21093032	22.350678243679845
tt3887868	22.128544574514734
tt1060269	22.10324910123272
tt15179648	22.065293009043472
tt7400462	22.02036791056055
tt7926112	22.001883775157047
tt18076786	21.95812859520454
tt4961164	21.9303108260282
tt12234698	21.87735008648564
tt13240610	21.87732078303079
tt8718578	21.81329261274322
tt9758542	21.705365319263116
tt21065316	21.69190619264592
tt8338828	21.690413724790545
tt4581420	21.68956790946911
tt4671296	21.665113854660653
tt15338648	21.625701986806693
tt2780908	21.586013045543154
tt15332976	21.57541647415407
tt3997022	21.57103994199139
tt17498564	21.554050889611823
tt5296996	21.531261279544974
tt15711350	21.506339918414923
tt2251838	21.486406042875732
tt9746718| 21.47634898932504|</td>
<td>('tt15405486', 0.00013307137120997648),
('tt15711350', 0.000133007817344747),
('tt15811964', 0.00013300478092529062),
('tt11663228', 0.0001328002484504905),
('tt11305742', 0.00013267435665719978),
('tt7926112', 0.00013236508358813035),
('tt2195857', 0.00013135698546262397),
('tt15249688', 0.00013077714965740019),
('tt11430948', 0.00013072326587841525),
('tt1060269', 0.00013019565243072356),
('tt11859296', 0.0001299616393925656),
('tt2173216', 0.0001298774535252642),
('tt12799846', 0.00012986265051332464),
('tt3529820', 0.00012949985780224331),
('tt14110888', 0.000129375655196421),
('tt7768574', 0.00012831774412177909),
('tt5371912', 0.00012830723852193932),
('tt2759098', 0.0001276476413287594),
('tt11834062', 0.00012756058070685473),
('tt8718578', 0.00012754594121979913),
('tt11465042', 0.00012712469013857882),
('tt10238724', 0.00012672607452367173),
('tt14348594', 0.00012654535392507473),
('tt14213908', 0.0001261555867273155),
('tt9041734', 0.000125962584541368),
('tt7462288', 0.00012497447202332093),
('tt7353622', 0.00012487711567118546),
('tt6338856', 0.0001245386080605285),
('tt6421868', 0.00012438306530595253),
('tt15357686', 0.00012369306811412152),
('tt11110938', 0.00012353593364615),
('tt18344978', 0.00012313582067016333),
('tt6913244', 0.0001230647505777476),
('tt12439942', 0.0001230368842366945),
('tt10928904', 0.00012286814121742007),
('tt11433936', 0.00012281344971969512),
('tt8338828', 0.00012265061919455481),
('tt21141138', 0.00012251845736083023),
('tt21093032', 0.00012251845736083023),
('tt12221126', 0.0001225103451324597),
('tt17498564', 0.00012244942100556744),
('tt15791002', 0.0001222951906535729),
('tt12234316', 0.0001221483718310262),
('tt5637536', 0.00012160258809290153),
('tt7400462', 0.00012140839726983328),
('tt14812120', 0.00012082365666408184),
('tt5567254', 0.000120516236718406),
('tt2780908', 0.00012044260982626298),
('tt2987506', 0.0001203188858427277),
('tt21572042', 0.00012020472097584247),
('tt9758542', 0.00012003377436553973),
('tt3997022', 0.00012000233373388862),
('tt10698588', 0.00011981492394286753),
('tt15909198', 0.00011963332018800064),
('tt9104736', 0.00011960773648275855),
('tt7283088', 0.00011950308186760793),
('tt5844538', 0.0001194794712658443),
('tt4945764', 0.00011931048999553168),
('tt8380948', 0.0001187667301261541),
('tt13178134', 0.00011870442257366293),</td>
</tr>
</table>

|tt10698588|21.431397351385716| |tt12973010|21.387507544967505| | tt4945764|21.326328125854324| |tt15909198| 21.23949194785293| | tt2353826|21.236128402782604| |tt15718740| 21.21593618512127| | tt8513794| 21.14349430517431| |tt10426796|21.121418550984593| |tt13671814| 21.1211382792741| | tt3529820|20.992419722867464| |tt14926914|20.986217747244865| |tt15094806| 20.97379010983368| |tt13667522| 20.96377901876255| | tt2993466| 20.9377083944112| | tt7626484|20.915705442501835| | tt9587298| 20.89612994353931| | tt2220128|20.824122527598334| |tt12893990| 20.80736354523577| | tt7768574| 20.80386358397028| | tt1380991| 20.73891185257837| |tt11574652| 20.7313114372153| |tt15526930|20.724330813844592| |tt14110458| 20.70278041232994| |tt14297962|20.636311310645716| |tt13797216|20.613670054921307| | tt5844538|20.584137420790164| | tt6913244|20.577603410397668| |tt14812120|20.553194455494626| |tt16420296| 20.55182060437179| |tt21253918|20.519379345362562| |tt10619392| 20.51208216855083| | tt8126470|20.502328257700583| |tt10238724|20.466648158287782| | tt9041734|20.438267466614466| |tt12564936|20.432739744912563| | tt6338856| 20.41874992867695| |tt14278792| 20.41372476412358| |tt11433936|20.371059294422604| | tt8407968| 20.32894580043828| | tt4386312|20.262387125769585| |tt14883936| 20.24026076708109| |tt10675748| 20.21853838348604| | tt7409674|20.191237481096078| |tt12648356|20.188743122337524| | tt5567254| 20.14926568096996| | tt7941914|20.111408578045367| |tt13391364|20.106384572713903| |tt15141416| 20.0942348415298| |tt10844184|19.910727378861658| |tt20563912|19.892796615896824| | tt6568416| 19.85793701372727| |tt16960240|19.755528174251786| | tt4087574| 19.74529269422045| | tt5637536| 19.74403872237949| | tt8814132|19.740848717043278| | tt9470210| 19.73274337236455| |tt14830912| 19.72233969461327| |tt15372082|19.696218903867262| |tt18934038|19.635802302725757| |tt19363670|19.606833324134957|	('tt16960240', 0.00011869404627930687), ('tt15338648', 0.00011861403035457381), ('tt6794554', 0.0001185851484267583), ('tt15526930', 0.00011854948182282986), ('tt12338246', 0.0001185078524758827), ('tt15536618', 0.00011845441025820675), ('tt2993466', 0.00011845175979652422), ('tt2365971', 0.00011844060158715612), ('tt10833362', 0.00011836290131678957), ('tt5296996', 0.00011823480442934019), ('tt5943284', 0.00011821905263107269), ('tt20918100', 0.00011812280093592135), ('tt8513794', 0.00011805070970042524), ('tt18934038', 0.00011797769348793036), ('tt15094806', 0.00011780147581920633), ('tt15141416', 0.00011779289735022993), ('tt10844184', 0.0001177550206805754), ('tt8136688', 0.00011755987631315904), ('tt15372082', 0.00011737598289601712), ('tt9470210', 0.00011727120381022134), ('tt15332976', 0.00011713755635713408), ('tt12564936', 0.00011710535266753055), ('tt3831118', 0.00011705987191232662), ('tt4087574', 0.00011654516710914255), ('tt9002366', 0.0001165401920819159), ('tt13240610', 0.00011605443387357511), ('tt15718740', 0.00011583754749045889), ('tt7626484', 0.00011544357780497937), ('tt16420296', 0.00011487790497856725), ('tt9358480', 0.0001145720008948423), ('tt1380991', 0.00011443943464502904), ('tt2220128', 0.00011440151667994834), ('tt3887868', 0.00011437603626842582), ('tt4671296', 0.00011405696456705312), ('tt14297962', 0.00011392664089933865), ('tt12234698', 0.00011371049424509221), ('tt9587298', 0.00011292367599350748), ('tt10426796', 0.00011289648339134878), ('tt13667522', 0.0001128355239957733), ('tt14527750', 0.00011272235585422456), ('tt4961164', 0.00011269643253331046), ('tt21253918', 0.00011258751603169053), ('tt14926914', 0.00011252124294801732), ('tt21065316', 0.00011233811160850874), ('tt18076786', 0.00011216396297540455), ('tt4250412', 0.00011214550539816341), ('tt2251838', 0.0001121431680682873 6), ('tt10414144', 0.00011189798629367011), ('tt15179648', 0.00011151879643823607), ('tt14278792', 0.0001115047073719978), ('tt16155570', 0.00011145707765538519), ('tt5953552', 0.0001111446435541456 8), ('tt12648356', 0.00011106157299622679), ('tt11548228', 0.00011098227681803285), ('tt14315416', 0.00011068577713285487), ('tt5959644', 0.0001106631942406388), ('tt11687216', 0.00011056316526127471), ('tt10675748', 0.0001103648308996894), ('tt12973010', 0.00011022995579786148), ('tt10619392', 0.00011019306822926556),

|tt15409934| 19.59758438462392|	('tt7941914', 0.00011012220430727664),
|tt16404438|19.575393739745497|	('tt14147254', 0.00011012081762424958),
| tt2987506| 19.5686946352983|	('tt2353826', 0.00011002599906231072),
| tt8246646| 19.55734172034491|	('tt4063092', 0.00010999234179791515),
| tt5953552|19.555823981255077|	('tt8239164', 0.00010989861991056964),
|tt15169036| 19.54811088183142|	('tt7409674', 0.00010971060309760116),
| tt9104736|19.463426213544842|	('tt12976188', 0.00010936882100672871),
|tt14527750| 19.45901914703139|	('tt9140180', 0.00010920333090005848),
| tt6799552| 19.35647702073959|	('tt13797216', 0.00010884773240727594),
|tt14147254| 19.3460183604294|	('tt14110458', 0.00010878982736243294),
|tt20834182| 19.3296869709451|	('tt8721852', 0.00010876637071926394),
| tt5705652| 19.32880580591648|	('tt15163896', 0.00010875772927524218),
|tt18482470|19.313615311197044|	('tt11574652', 0.00010875507263783575),
| tt7283088| 19.30045837756502|	('tt13671814', 0.00010872946743239931),
|tt13178134|19.190539714583505|	('tt4561132', 0.00010858707999497325),
| tt9702134|19.125147412863893|	('tt18279118', 0.00010852067271374776),
|tt11076278|19.079011557345787|	('tt15169036', 0.00010841849850970592),
|tt14286576|19.073850433139526|	('tt11193246', 0.00010815547283857892),
| tt9002366|19.032916857668155|	('tt13192446', 0.00010814736117169434),
| tt8380948| 19.02580874569751|	('tt12893990', 0.00010814547822328862),
|tt12598744|19.007344169126952|	('tt9746718', 0.0001080307818362281),
|tt11687216| 18.98639263736237|	('tt21361444', 0.00010802174569352639),
| tt3907746| 18.9101333296492|	('tt10418818', 0.00010789396338874499),
|tt20864606|18.862614302407017|	('tt9418048', 0.00010788903652992832),
|tt13586204|18.843065668350125|	('tt14883936', 0.00010774083934675323),
| tt6794554|18.839631233858196|	('tt8126470', 0.00010718603505615519)]

Historic Approach 20 ITER, 200 Movies	Probabilistic Approach
id pagerank	id pagerank
|tt11916562| 39.60245884175969|	[('tt4745692', 0.0002163202093671938),
| tt4745692|39.442991818003655|	('tt3886736', 0.0002119182217079988),
| tt3886736| 37.80073131259984|	('tt11916562', 0.0002077988514825338),
| tt4968564| 36.19073002151844|	('tt5918332', 0.00018751721057302289),
|tt10765798| 35.37464543176991|	('tt10765798', 0.00018525163612327557),
| tt5918332| 34.77857872337126|	('tt4968564', 0.0001843011357755323),
|tt13114794| 33.77550647509047|	('tt13114794', 0.00017800891455590747),
|tt14132588| 31.41995527539344|	('tt21215814', 0.00016932777192667694),
| tt9512094|31.087462854691783|	('tt12399882', 0.00016801499272454175),
|tt21215814|30.980978912141293|	('tt11776266', 0.00016505671892582982),
|tt13950392|30.937774993113365|	('tt9512094', 0.0001607731867347449),
|tt14416808| 30.80134662924967|	('tt6496598', 0.0001603969706365523),
|tt12399882|30.545687692929878|	('tt13950392', 0.00015984858831305464),
|tt12126628|30.103351732781544|	('tt14132588', 0.0001592065264466762),
|tt10921118|29.725833473910278|	('tt12126628', 0.00015743126715840616),
|tt12709774| 29.71092271080792|	('tt10921118', 0.00015666348459495585),
| tt3643390|29.494070665436922|	('tt14416808', 0.00015602474027694347),
|tt11075800|29.485127876320423|	('tt11075800', 0.0001542162388808778),
|tt11776266| 29.45431171448077|	('tt3643390', 0.0001540468186092864),
|tt11416582|29.318910908517562|	('tt12709774', 0.00015383212108394708),

|tt19387356| 28.83363581194299|	('tt13498286', 0.00015322735672170244),
|tt10670022|28.681638793371878|	('tt13904218', 0.00015264265665299772),
|tt14069222|28.116447466284278|	('tt19387356', 0.00015242689782367857),
|tt14182380| 27.86383142403571|	('tt11416582', 0.00015084335228583994),
|tt15847776| 27.22761202399111|	('tt14182380', 0.0001499979232272674),
|tt17278774| 27.12057222274021|	('tt10670022', 0.00014997397108241917),
|tt14997306|27.113507849174532|	('tt14069222', 0.0001476860273394416),
|tt13904218|27.097293529291054|	('tt2397595', 0.0001463559637863334),
|tt12658830|27.029154194197076|	('tt16450020', 0.00014632526022390662),
| tt6496598| 26.75855479571765|	('tt5167166', 0.0001457807103193064),
|tt12498996| 26.66925963535965|	('tt12498996', 0.0001452195510055958),
|tt17155092|26.613561565173868|	('tt17155092', 0.00014433345974199362),
| tt2397595|26.385629033396754|	('tt16048054', 0.00014328561892115702),
|tt12752880| 26.32560331309736|	('tt1174971', 0.0001431154487860039),
|tt21239560| 26.02203480066064|	('tt15847776', 0.0001417591495341676),
|tt13542474| 25.85856515646123|	('tt17278774', 0.00014095415158886493),
|tt13498286|25.791252392444637|	('tt16760028', 0.00014062225218721013),
|tt15825262| 25.75328979333861|	('tt4167532', 0.00013928131398403983),
| tt8435252|25.635898159855746|	('tt12658830', 0.0001386575390296042),
|tt11305742|25.450994699110897|	('tt4390908', 0.00013850764210050512),
|tt12623450|25.390726631619273|	('tt8435252', 0.00013838878990371257),
|tt15405486|25.247275691266598|	('tt12752880', 0.00013818419964414805),
| tt5167166| 25.18855998830985|	('tt13542474', 0.00013809862990901342),
| tt1174971|25.146505894114252|	('tt12623450', 0.0001378899864140864),
|tt16048054| 25.11309483013625|	('tt4581992', 0.000137776931016182),
|tt12799846| 25.09433517645111|	('tt8586578', 0.000137614308845436),
| tt9169200|24.893568160013853|	('tt3143458', 0.00013640961950381222),
| tt4390908|24.823939731434944|	('tt13993652', 0.00013564855108873778),
|tt16760028|24.808228593089893|	('tt21239560', 0.00013557282449204187),
|tt15811964|24.802365836835165|	('tt14997306', 0.00013543734876016316),
|tt12768728| 24.71901262688493|	('tt15825262', 0.00013496093112624262),
| tt2195857| 24.65106988293704|	('tt4581420', 0.00013467450739077476),
|tt11834062|24.626912346595375|	('tt12768728', 0.0001344972332361274),
|tt11430948|24.511181520960026|	('tt9169200', 0.00013424480764509723),
| tt4167532|24.345271484651207|	('tt15405486', 0.00013307137120997648),
| tt8586578|24.318209543919252|	('tt15711350', 0.000133007817344747),
|tt12439942| 24.275716761322|	('tt15811964', 0.00013300478092529062),
|tt11859296|24.201742377383813|	('tt11663228', 0.0001328002484504905),
| tt4581992| 24.00163248222558|	('tt11305742', 0.00013267435665719978),
| tt2759098|23.940629670747438|	('tt7926112', 0.00013236508358813035),
|tt16450020|23.894127765723233|	('tt2195857', 0.00013135698546262397),
|tt15357686| 23.82859632126881|	('tt15249688', 0.00013077714965740019),
| tt3831118|23.798984343805714|	('tt11430948', 0.00013072326587841525),
| tt7462288| 23.73601060840957|	('tt1060269', 0.00013019565243072356),
|tt10833362|23.578935329559375|	('tt11859296', 0.0001299616393925656),
|tt14110888|23.563986945281837|	('tt2173216', 0.0001298774535252642),
|tt18344978|23.475950693883235|	('tt12799846', 0.00012986265051332464),
|tt15791002|23.342743069692492|	('tt3529820', 0.00012949985780224331),
|tt14348594|23.278978747936417|	('tt14110888', 0.000129375655196421),
|tt15536618|23.211521731439838|	('tt7768574', 0.00012831774412177909),
| tt2173216|23.168657248651286|	('tt5371912', 0.00012830723852193932),
| tt5371912|23.118330465654733|	('tt2759098', 0.0001276476413287594),
|tt21572042|23.101302133606797|	('tt11834062', 0.00012756058070685473),
|tt10928904|22.979324361437907|	('tt8718578', 0.00012754594121979913),
|tt15249688| 22.95723290722756|	('tt11465042', 0.00012712469013857882),
|tt11663228| 22.88556536760316|	('tt10238724', 0.00012672607452367173),
| tt3143458|22.871672614293065|	('tt14348594', 0.00012654535392507473),
|tt20918100|22.867031988128648|	('tt14213908', 0.0001261555867273155),
| tt6421868|22.849902258012524|	('tt9041734', 0.000125962584541368),
|tt12221126|22.835680837521704|	('tt7462288', 0.00012497447202332093),

|tt13993652|22.742166291892698|	('tt7353622', 0.00012487711567118546),
|tt11465042|22.731711795665063|	('tt6338856', 0.0001245386080605285),
|tt12234316|22.715206799633574|	('tt6421868', 0.00012438306530595253),
|tt11110938|22.697475347813658|	('tt15357686', 0.00012369306811412152),
| tt2365971|22.674755666156006|	('tt11110938', 0.00012353593364615),
|tt14213908| 22.66076434844518|	('tt18344978', 0.00012313582067016333),
| tt7353622| 22.47783029293113|	('tt6913244', 0.000123064750577476),
|tt12338246|22.408920266227483|	('tt12439942', 0.0001230368842366945),
|tt21141138|22.350678243679848|	('tt10928904', 0.00012286814121742007),
|tt21093032|22.350678243679845|	('tt11433936', 0.00012281344971969512),
| tt3887868|22.128544574514734|	('tt8338828', 0.00012265061919455481),
| tt1060269| 22.10324910123272|	('tt21141138', 0.00012251845736083023),
|tt15179648|22.065293009043472|	('tt21093032', 0.00012251845736083023),
| tt7400462| 22.02036791056055|	('tt12221126', 0.0001225103451324597),
| tt7926112|22.001883775157047|	('tt17498564', 0.00012244942100556744),
|tt18076786| 21.95812859520454|	('tt15791002', 0.0001222951906535729),
| tt4961164| 21.9303108260282|	('tt12234316', 0.0001221483718310262),
|tt12234698| 21.87735008648564|	('tt5637536', 0.00012160258809290153),
|tt13240610| 21.87732078303079|	('tt7400462', 0.00012140839726983328),
| tt8718578| 21.81329261274322|	('tt14812120', 0.00012082365666408184),
| tt9758542|21.705365319263116|	('tt5567254', 0.000120516236718406),
|tt21065316| 21.69190619264592|	('tt2780908', 0.00012044260982626298),
| tt8338828|21.690413724790545|	('tt2987506', 0.0001203188858427277),
| tt4581420| 21.68956790946911|	('tt21572042', 0.00012020472097584247),
| tt4671296|21.665113854660653|	('tt9758542', 0.00012003377436553973),
|tt15338648|21.625701986806693|	('tt3997022', 0.00012000233373388862),
| tt2780908|21.586013045543154|	('tt10698588', 0.00011981492394286753),
|tt15332976| 21.57541647415407|	('tt15909198', 0.00011963332018800064),
| tt3997022| 21.57103994199139|	('tt9104736', 0.0001196077364827585),
|tt17498564|21.554050889611823|	('tt7283088', 0.00011950308186760793),
| tt5296996|21.531261279544974|	('tt5844538', 0.0001194794712658443),
|tt15711350|21.506339918414923|	('tt4945764', 0.00011931048999553168),
| tt2251838|21.486406042875732|	('tt8380948', 0.0001187667301261541),
| tt9746718| 21.47634898932504|	('tt13178134', 0.00011870442257366293),
|tt10698588|21.431397351385716|	('tt16960240', 0.00011869404627930687),
|tt12973010|21.387507544967505|	('tt15338648', 0.00011861403035457381),
| tt4945764|21.326328125854324|	('tt6794554', 0.0001185851484267583),
|tt15909198| 21.23949194785293|	('tt15526930', 0.00011854948182282986),
| tt2353826|21.236128402782604|	('tt12338246', 0.0001185078524758827),
|tt15718740| 21.21593618512127|	('tt15536618', 0.00011845441025820675),
| tt8513794| 21.14349430517431|	('tt2993466', 0.00011845175979652422),
|tt10426796|21.121418550984593|	('tt2365971', 0.0001184406015871561 2),
|tt13671814| 21.1211382792741|	('tt10833362', 0.00011836290131678957),
| tt3529820|20.992419722867464|	('tt5296996', 0.00011823480442934019),
|tt14926914|20.986217747244865|	('tt5943284', 0.00011821905263107269),
|tt15094806| 20.97379010983368|	('tt20918100', 0.00011812280093592135),
|tt13667522| 20.96377901876255|	('tt8513794', 0.00011805070970042524),
| tt2993466| 20.9377083944112|	('tt18934038', 0.00011797769348793036),
| tt7626484|20.915705442501835|	('tt15094806', 0.0001178014758192063 3),
| tt9587298| 20.89612994353931|	('tt15141416', 0.00011779289735022993),
| tt2220128|20.824122527598334|	('tt10844184', 0.0001177550206805754),
|tt12893990| 20.80736354523577|	('tt8136688', 0.00011755987631315904),
| tt7768574| 20.80386358397028|	('tt15372082', 0.00011737598289601712),
| tt1380991| 20.73891185257837|	('tt9470210', 0.00011727120381022134),
|tt11574652| 20.7313114372153|	('tt15332976', 0.00011713755635713408),
|tt15526930|20.724330813844592|	('tt12564936', 0.00011710535266753055),
|tt14110458| 20.70278041232994|	('tt3831118', 0.00011705987191232662),
|tt14297962|20.636311310645716|	('tt4087574', 0.00011654516710914255),
|tt13797216|20.613670054921307|	('tt9002366', 0.0001165401920819159),
| tt5844538|20.584137420790164|	('tt13240610', 0.00011605443387357511),

<table>
<tr><td>| tt6913244|20.577603410397668|</td><td>('tt15718740', 0.00011583754749045889),</td></tr>
<tr><td>|tt14812120|20.553194455494626|</td><td>('tt7626484', 0.00011544357780497937),</td></tr>
<tr><td>|tt16420296| 20.55182060437179|</td><td>('tt16420296', 0.00011487790497856725),</td></tr>
<tr><td>|tt21253918|20.519379345362562|</td><td>('tt9358480', 0.0001145720008948423),</td></tr>
<tr><td>|tt10619392| 20.51208216855083|</td><td>('tt1380991', 0.00011443943464502904),</td></tr>
<tr><td>| tt8126470|20.502328257700583|</td><td>('tt2220128', 0.00011440151667994834),</td></tr>
<tr><td>|tt10238724|20.466648158287782|</td><td>('tt3887868', 0.00011437603626842582),</td></tr>
<tr><td>| tt9041734|20.438267466614466|</td><td>('tt4671296', 0.00011405696456705312),</td></tr>
<tr><td>|tt12564936|20.432739744912563|</td><td>('tt14297962', 0.00011392664089933865),</td></tr>
<tr><td>| tt6338856| 20.41874992867695|</td><td>('tt12234698', 0.00011371049424509221),</td></tr>
<tr><td>|tt14278792| 20.41372476412358|</td><td>('tt9587298', 0.00011292367599350748),</td></tr>
<tr><td>|tt11433936|20.371059294422604|</td><td>('tt10426796', 0.00011289648339134878),</td></tr>
<tr><td>| tt8407968| 20.32894580043828|</td><td>('tt13667522', 0.0001128355239957733),</td></tr>
<tr><td>| tt4386312|20.262387125769585|</td><td>('tt14527750', 0.00011272235585422456),</td></tr>
<tr><td>|tt14883936| 20.24026076708109|</td><td>('tt4961164', 0.00011269643253331046),</td></tr>
<tr><td>|tt10675748| 20.21853838348604|</td><td>('tt21253918', 0.00011258751603169053),</td></tr>
<tr><td>| tt7409674|20.191237481096078|</td><td>('tt14926914', 0.00011252124294801732),</td></tr>
<tr><td>|tt12648356|20.188743122337524|</td><td>('tt21065316', 0.00011233811160850874),</td></tr>
<tr><td>| tt5567254| 20.14926568096996|</td><td>('tt18076786', 0.00011216396297540455),</td></tr>
<tr><td>| tt7941914|20.111408578045367|</td><td>('tt4250412', 0.00011214550539816341),</td></tr>
<tr><td>|tt13391364|20.106384572713903|</td><td>('tt2251838', 0.00011214316806828736),</td></tr>
<tr><td>|tt15141416| 20.0942348415298|</td><td>('tt10414144', 0.00011189798629367011),</td></tr>
<tr><td>|tt10844184|19.910727378861658|</td><td>('tt15179648', 0.00011151879643823607),</td></tr>
<tr><td>|tt20563912|19.892796615896824|</td><td>('tt14278792', 0.0001115047073719978),</td></tr>
<tr><td>| tt6568416| 19.85793701372727|</td><td>('tt16155570', 0.00011145707765538519),</td></tr>
<tr><td>|tt16960240|19.755528174251786|</td><td>('tt5953552', 0.00011114464355414568),</td></tr>
<tr><td>| tt4087574| 19.74529269422045|</td><td>('tt12648356', 0.00011106157299622679),</td></tr>
<tr><td>| tt5637536| 19.74403872237949|</td><td>('tt11548228', 0.00011098227681803285),</td></tr>
<tr><td>| tt8814132|19.740848717043278|</td><td>('tt14315416', 0.00011068577713285487),</td></tr>
<tr><td>| tt9470210| 19.73274337236455|</td><td>('tt5959644', 0.0001106631942406388),</td></tr>
<tr><td>|tt14830912| 19.72233969461327|</td><td>('tt11687216', 0.00011056316526127471),</td></tr>
<tr><td>|tt15372082|19.696218903867262|</td><td>('tt10675748', 0.0001103648308996894),</td></tr>
<tr><td>|tt18934038|19.635802302725757|</td><td>('tt12973010', 0.00011022995579786148),</td></tr>
<tr><td>|tt19363670|19.606833324134957|</td><td>('tt10619392', 0.00011019306822926556),</td></tr>
<tr><td>|tt15409934| 19.59758438462392|</td><td>('tt7941914', 0.00011012220430727664),</td></tr>
<tr><td>|tt16404438|19.575393739745497|</td><td>('tt14147254', 0.00011012081762424958),</td></tr>
<tr><td>| tt2987506| 19.5686946352983|</td><td>('tt2353826', 0.00011002599906231072),</td></tr>
<tr><td>| tt8246646| 19.55734172034491|</td><td>('tt4063092', 0.00010999234179791515),</td></tr>
<tr><td>| tt5953552|19.555823981255077|</td><td>('tt8239164', 0.00010989861991056964),</td></tr>
<tr><td>|tt15169036| 19.54811088183142|</td><td>('tt7409674', 0.00010971060309760116),</td></tr>
<tr><td>| tt9104736|19.463426213544842|</td><td>('tt12976188', 0.00010936882100672871),</td></tr>
<tr><td>|tt14527750| 19.45901914703139|</td><td>('tt9140180', 0.00010920333090005848),</td></tr>
<tr><td>| tt6799552| 19.35647702073959|</td><td>('tt13797216', 0.00010884773240727594),</td></tr>
<tr><td>|tt14147254| 19.3460183604294|</td><td>('tt14110458', 0.00010878982736243294),</td></tr>
<tr><td>|tt20834182| 19.3296869709451|</td><td>('tt8721852', 0.00010876637071926394),</td></tr>
<tr><td>| tt5705652| 19.32880580591648|</td><td>('tt15163896', 0.00010875772927524218),</td></tr>
<tr><td>|tt18482470|19.313615311197044|</td><td>('tt11574652', 0.00010875507263783575),</td></tr>
<tr><td>| tt7283088| 19.30045837756502|</td><td>('tt13671814', 0.00010872946743239931),</td></tr>
<tr><td>|tt13178134|19.190539714583505|</td><td>('tt4561132', 0.00010858707999497325),</td></tr>
<tr><td>| tt9702134|19.125147412863893|</td><td>('tt18279118', 0.00010852067271374776),</td></tr>
<tr><td>|tt11076278|19.079011557345787|</td><td>('tt15169036', 0.00010841849850970592),</td></tr>
<tr><td>|tt14286576|19.073850433139526|</td><td>('tt11193246', 0.00010815547283857892),</td></tr>
<tr><td>| tt9002366|19.032916857668155|</td><td>('tt13192446', 0.00010814736117169434),</td></tr>
<tr><td>| tt8380948| 19.02580874569751|</td><td>('tt12893990', 0.00010814547822328862),</td></tr>
<tr><td>|tt12598744|19.007344169126952|</td><td>('tt9746718', 0.0001080307818362281),</td></tr>
<tr><td>|tt11687216| 18.98639263736237|</td><td>('tt21361444', 0.00010802174569352639),</td></tr>
<tr><td>| tt3907746| 18.9101333296492|</td><td>('tt10418818', 0.00010789396338874499),</td></tr>
<tr><td>|tt20864606|18.862614302407017|</td><td>('tt9418048', 0.00010788903652992832),</td></tr>
<tr><td>|tt13586204|18.843065668350125|</td><td>('tt14883936', 0.00010774083934675323),</td></tr>
<tr><td>| tt6794554|18.839631233858196|</td><td>('tt8126470', 0.00010718603505615519)]</td></tr>
</table>