Steps Followed To Complete The Assignment

Initial Setup

I installed virtual machine with ubuntu as operating system. Then, I installed Solr along with downloading FOX_NEWS data in parallel.

Lucene Result Setup

After that, I made necessary changes in "managed-schema.xml" file and started the Solr on port 8989. Then, executed command to get page rank of the HTML files for my core "myexample". Then after successful execution of index core was reloaded by using core admin from solr user interface. After that, I installed solr-php-client and using it, created user interface to take query input from user and display top 10 result. This code was written in PHP in file name named "user.php". In this, while querying additional parameters were not used. For each of the query, top 10 url results were noted in below table. Each item of list displays document title, id, description and link. Link and title are clickable and would direct to corresponding html page when clicked. For documentation flow, query "Bitcoin Blockchain" screen shot was taken as per the requirement.

PageRank Result Setup

In order to implement PageRank, JSOUP was downloaded. I wrote a java code to create outgoing edges from a html page using JSOUP. This file was named "external_url.txt". ""external_url.txt" served as an input to python code where a graph was generated using "NetworkX" library. On this graph, pagerank of each node was calculated with parameters as "alpha=0.85, personalization=None, max_iter=30, tol=1e-06, nstart=None, weight='weight',dangling=None". Once the pagerank was calculated successfully, each node(document) was saved as "documentId=documentRank" in "external_PageRankFile.txt". Now, "user.php" file was changed to include additional parameter as "sort: PageRankFile desc" and "external_PageRankFile.txt" was placed in "data" folder under "myexample" core after updating "managed-schema.xml" and "solrconfig.xml" files. This completed the setup for PageRank algorithm. Again from solr user interface, core was refreshed. Once again, each query was executed and top 10 urls were noted in below table. For documentation flow, query "Bitcoin Blockchain" screen shot was taken as per the requirement.

Overlapping Graph

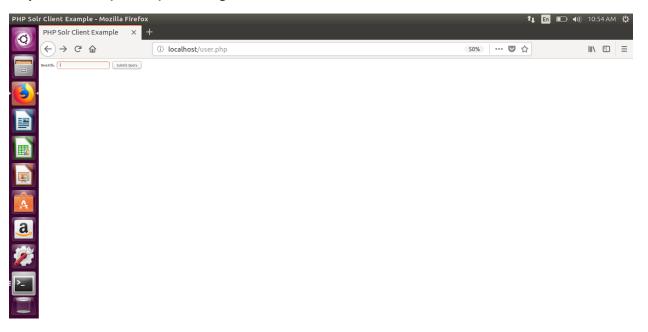
Once urls were recorded in table for default and PageRank, overlapping graph was plotted as the final step for the assignment. For a query, if the difference between url result of lucene(default) and PageRank had only difference in domain or protocol then I considered it not as an overlap.

Explanation Regarding Why Some Pages Have Higher Page Rank

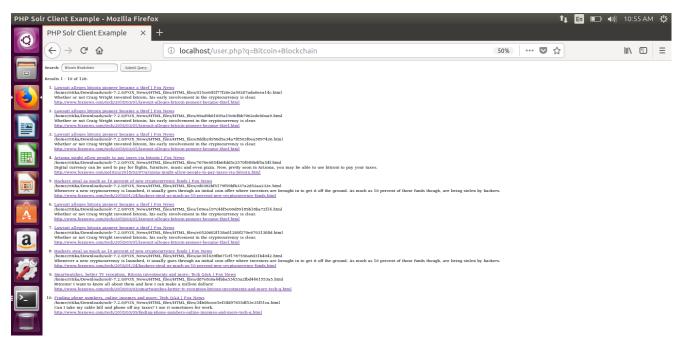
PageRank is calculated on the basis of number of incoming links to a page. Let's say we have page1 with no incoming links and page2 with 1 incoming links and we only have 2 link. In this case, page2 will have more page rank than page1. The reason why "http://www.foxnews.com" is ranked first in query 1 is that it is homepage. Therefore, almost all other website would have link to it. Hence it's PageRank is highest.

Document Flow Of Single Query

Step 1.1: Lucene(Default) Initial Page



Step 1.2: Lucene(Default) Query Result for Bitcoin Blockchain

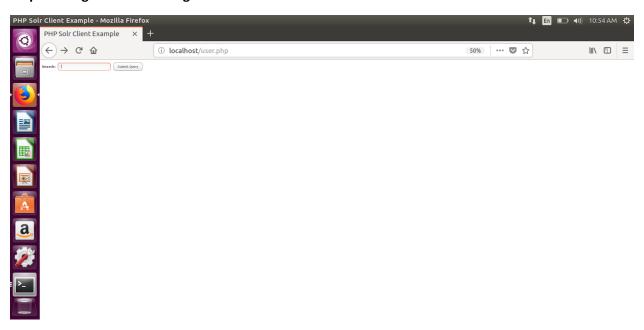


Step 1.3: Lucene(Default) Link Page Upon Clicking first title

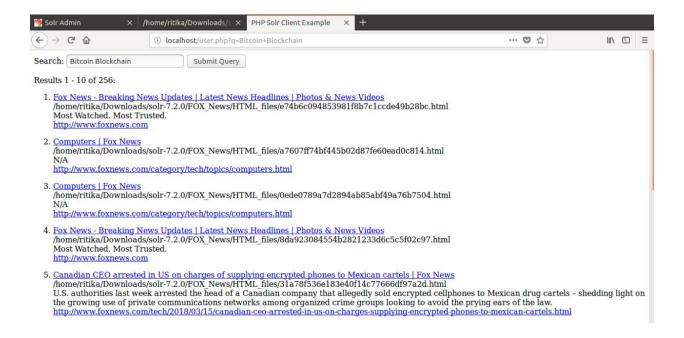


Add "additionalParameter" to query input and proceed to following steps.

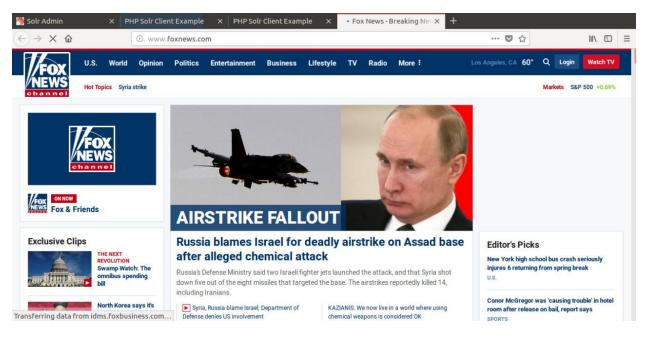
Step 2.1: PageRank Initial Page



Step 2.2: PageRank Query Result for Bitcoin Blockchain



Step 2.3: PageRank Link Page Upon Clicking first title



URL Results For Query

Query 1: Bitcoin Blockchain

| Ra | Lucene (Default) Result | PageRank Result |
|----|---|---|
| nk | | |
| 1 | http://www.foxnews.com/tech/2018/03/01 | http://www.foxnews.com |
| | /lawsuit-alleges-bitcoin-pioneer-became- | |
| | <u>thief.html</u> | |
| 2 | http://www.foxnews.com/tech/2018/03/01 | http://www.foxnews.com/category/tech/topics |
| | <u>/lawsuit-alleges-bitcoin-pioneer-became-</u> | /computers.html |
| | <u>thief.html</u> | |
| 3 | http://www.foxnews.com/tech/2018/03/01 | http://www.foxnews.com/category/tech/topics |
| | <u>/lawsuit-alleges-bitcoin-pioneer-became-</u> | /computers.html |
| | <u>thief.html</u> | |
| 4 | http://www.foxnews.com/politics/2018/02 | http://www.foxnews.com |
| | /07/arizona-might-allow-people-to-pay- | |
| | taxes-via-bitcoin.html | |
| 5 | http://www.foxnews.com/tech/2018/01/24 | http://www.foxnews.com/tech/2018/03/15/can |
| | /hackers-steal-as-much-as-10-percent-new- | adian-ceo-arrested-in-us-on-charges-supplying- |
| | <u>cryptocurrency-funds.html</u> | encrypted-phones-to-mexican-cartels.html |
| 6 | http://www.foxnews.com/tech/2018/03/01 | http://www.foxnews.com/tech/2018/03/02/eq |
| | <u>/lawsuit-alleges-bitcoin-pioneer-became-</u> | <u>uifax-breach-worsens-sparking-more-concerns-</u> |
| | thief.html | <u>about-consumers-data.html</u> |
| 7 | http://www.foxnews.com/tech/2018/03/01 | http://www.foxnews.com/tech/2018/03/02/big |
| | <u>/lawsuit-alleges-bitcoin-pioneer-became-</u> | gest-ddos-attack-on-record-hits-github.html |
| | thief.html | |
| 8 | http://www.foxnews.com/tech/2018/01/24 | http://www.foxnews.com/tech/2018/02/22/hac |
| | /hackers-steal-as-much-as-10-percent-new- | kers-hijack-tesla-system-in-bold-cryptocurrency- |
| | <u>cryptocurrency-funds.html</u> | attack.html |
| 9 | http://www.foxnews.com/tech/2018/03/03 | http://www.foxnews.com/tech/2018/02/04/tec |
| | /smartwatches-better-tv-reception-bitcoin- | h-q-dueling-assistants-outsmarting-hackers- |
| | <u>investments-and-more-tech-q.html</u> | blocking-texts-and-more.html |
| 10 | http://www.foxnews.com/tech/2018/03/09 | http://www.foxnews.com/category/tech/topics |
| | /finding-phone-numbers-online-incomes- | /cybercrime.html |
| | and-more-tech-q.html | |

Query 2: Stephen Hawking

| Ra nk | Lucene (Default) Result | PageRank Result |
|----------|---|---|
| 1 | http://www.foxnews.com/science/slideshow/ 2018/03/14/look-back-on-life-renowned- physicist-stephen-hawking.html | http://www.foxnews.com/science.html |
| 2 | http://www.foxnews.com/science/slideshow/ 2018/03/14/look-back-on-life-renowned- physicist-stephen-hawking.html | http://www.foxnews.com/opinion.html |
| 3 | http://www.foxnews.com/science/2018/03/16 /how-stephen-hawking-transformed-our- understanding-black-holes.html | http://www.foxnews.com/entertainment.ht ml |

| 4 | http://www.foxnews.com/science/2018/03/16/how-stephen-hawking-transformed-our-understanding-black-holes.html | http://www.foxnews.com/category/science/ natural-science.html |
|----|--|--|
| 5 | http://www.foxnews.com/science/2018/03/14 /stephen-hawking-unlocked-universe- possibilities-nasa-says.html | http://www.foxnews.com/category/science/ air-and-space.html |
| 6 | http://www.foxnews.com/science/2018/03/14 /stephen-hawking-unlocked-universe- possibilities-nasa-says.html | http://www.foxnews.com/category/world/global-economy.html |
| 7 | http://www.foxnews.com/science/2018/03/14 /stephen-hawking-famed-physicist-dead-at- 76.html | http://www.foxnews.com/category/us/personal-freedoms.html |
| 8 | http://www.foxnews.com/science/2018/03/14 /stephen-hawking-famed-physicist-dead-at- 76.html | http://www.foxnews.com/category/tech/topics/security.html |
| 9 | http://www.foxnews.com/tech/2018/03/14/st ephen-hawking-tech-execs-remember-legacy- famed-physicist.html | http://www.foxnews.com/science.html |
| 10 | http://www.foxnews.com/tech/2018/03/14/st ephen-hawking-tech-execs-remember-legacy- famed-physicist.html | http://www.foxnews.com/entertainment.ht ml |

Query 3: Falcon Heavy

| Ran | Lucene (Default) Result | PageRank Result |
|-----|--|---|
| k | | |
| 1 | http://www.foxnews.com/science/2018/03/12/how-spacex-falcon-heavy-rocket- | http://www.foxnews.com/weather/your- weather/index.html |
| | compares-to-elon-musks-other-powerful- spacecraft.html | wedther/macx.ntm |
| 2 | http://www.foxnews.com/science/2018/03/12/how-spacex-falcon-heavy-rocket-compares-to-elon-musks-other-powerful-spacecraft.html | http://www.foxnews.com/category/world/disas ters.html |
| 3 | http://www.foxnews.com/science/2018/02/06/spacex-falcon-heavy-rocket-launches-successfully.html | http://www.foxnews.com/category/tech/topics/security.html |
| 4 | http://www.foxnews.com/science/2018/03/12/how-spacex-falcon-heavy-rocket-compares-to-elon-musks-other-powerful-spacecraft.html | http://www.foxnews.com/category/politics/elections.html |
| 5 | http://www.foxnews.com/science/2018/02/06/spacex-falcon-heavy-rocket-launches-successfully.html | http://www.foxnews.com/category/health/men tal-health.html |
| 6 | http://www.foxnews.com/science/2018/03/12/how-spacex-falcon-heavy-rocket- | http://www.foxnews.com/weather/index.html |

| | compares-to-elon-musks-other-powerful- | |
|----|--|---|
| | spacecraft.html | |
| 7 | http://www.foxnews.com/science/2018/03 | http://www.foxnews.com/weather/index.html |
| | /12/how-spacex-falcon-heavy-rocket- | |
| | compares-to-elon-musks-other-powerful- | |
| | spacecraft.html | |
| 8 | http://www.foxnews.com/science/2018/02 | http://www.foxnews.com/category/tech/topics/ |
| | /06/spacex-falcon-heavy-rocket-launches- | security.html |
| | successfully.html | |
| 9 | http://www.foxnews.com/science/2018/02 | http://www.foxnews.com/category/politics/elec |
| | /06/spacex-falcon-heavy-launch-preview- | tions.html |
| | what-when-and-why.html | |
| 10 | http://www.foxnews.com/science/2018/02 | http://www.foxnews.com/category/health/men |
| | /07/elon-musk-has-launched-falcon-heavy- | tal-health.html |
| | but-how-long-does-it-take-to-get-to- | |
| | mars.html | |

Query 4: Syrian War

| Ra nk | Lucene (Default) Result | PageRank Result |
|----------|---|--|
| 1 | http://www.foxnews.com/world/2018 /03/15/syria-marks-7-years-war- | http://www.foxnews.com/world.html |
| 2 | hundreds-leave-besieged-enclave.html http://www.foxnews.com/world/2018 /03/15/syria-marks-7-years-war- hundreds-leave-besieged-enclave.html | http://www.foxnews.com/us.html |
| 3 | http://www.foxnews.com/world/2018/03/15/syria-marks-7-years-war-hundreds-leave-besieged-enclave.html | http://www.foxnews.com/science.html |
| 4 | http://www.foxnews.com/world/2018 /03/15/syria-marks-7-years-war- hundreds-leave-besieged-enclave.html | http://www.foxnews.com/opinion.html |
| 5 | http://www.foxnews.com/world/2018 /03/15/syria-marks-7-years-war- hundreds-leave-besieged-enclave.html | http://www.foxnews.com/shows.html |
| 6 | http://www.foxnews.com/politics/201 7/03/10/new-trump-travel-ban-cant- be-enforced-on-syrian-family-trying- to-flee-to-wisconsin-judge-says.html | http://www.foxnews.com/weather/your-weather/index.html |
| 7 | http://www.foxnews.com/opinion/20 18/03/16/rep-adam-kinzinger-it-s-for- us-to-lead-on-syria-failure-to-act-is- failure-by-all-us.html | http://www.foxnews.com/auto.html |
| 8 | http://www.foxnews.com/opinion/20 18/03/16/rep-adam-kinzinger-it-s-for- | http://www.foxnews.com/category/science/archaeology.html |

| | us-to-lead-on-syria-failure-to-act-is- | |
|----|--|--|
| | failure-by-all-us.html | |
| 9 | http://www.foxnews.com/opinion/20 | http://www.foxnews.com/category/science/archaeol |
| | 18/03/16/rep-adam-kinzinger-it-s-for- | ogy/dinosaurs.html |
| | us-to-lead-on-syria-failure-to-act-is- | |
| | failure-by-all-us.html | |
| 10 | http://www.foxnews.com/world/2017 | http://www.foxnews.com/category/science/planet- |
| | /04/08/in-istanbul-little-syria- | earth.html |
| | refugees-want-more-from-us.html | |

Query 5: LA Rams

| Ran | Lucene (Default) Result | PageRank Result |
|-----|---|---|
| k | , | |
| 1 | http://www.foxnews.com/sports/2017/01/ | http://www.foxnews.com |
| | 25/jim-kelly-draws-on-experience-in- | |
| | campaign-against-cancer.html | |
| 2 | http://www.foxnews.com/us/2018/02/04/ | http://www.foxnews.com/tech.html |
| | super-bowl-protesters-to-take-knee-in- | |
| | minneapolis-with-rally-against-police- | |
| | <u>brutality.html</u> | |
| 3 | http://www.foxnews.com/us/2018/02/04/ | http://www.foxnews.com/lifestyle.html |
| | super-bowl-protesters-to-take-knee-in- | |
| | minneapolis-with-rally-against-police- | |
| | <u>brutality.html</u> | |
| 4 | http://www.foxnews.com/sports/2018/01/ | http://www.foxnews.com/category/style-and- |
| | 24/rams-add-former-ucla-offensive- | <u>beauty.html</u> |
| | coordinator-jedd-fisch-to-staff.html | |
| 5 | http://www.foxnews.com/sports/2018/01/ | http://www.foxnews.com/category/politics/elec |
| | 24/rams-add-former-ucla-offensive- | tions.html |
| | coordinator-jedd-fisch-to-staff.html | |
| 6 | http://www.foxnews.com/sports/2018/01/ | http://www.foxnews.com/tech.html |
| | 24/rams-add-former-ucla-offensive- | |
| | coordinator-jedd-fisch-to-staff.html | |
| 7 | http://www.foxnews.com/sports/2018/01/ | http://www.foxnews.com/lifestyle.html |
| | 24/rams-add-former-ucla-offensive- | |
| | coordinator-jedd-fisch-to-staff.html | |
| 8 | http://www.foxnews.com/sports/2018/01/ | http://www.foxnews.com/category/style-and- |
| | 24/rams-add-former-ucla-offensive- | <u>beauty.html</u> |
| | coordinator-jedd-fisch-to-staff.html | |
| 9 | http://www.foxnews.com/auto/2018/01/2 | http://www.foxnews.com/category/politics/elec |
| | 3/tesla-allegedly-on-autopilot-rams-into- | <u>tions.html</u> |
| | back-fire-truck.html | |
| 10 | http://www.foxnews.com/us/2018/02/04/ | http://www.foxnews.com/shows/cavuto- |
| | super-bowl-protesters-to-take-knee-in- | <u>live.html</u> |

| minneapolis-with-rally-against-police- |
|--|
| <u>brutality.html</u> |

Query 6: Black Panther

| Ran | Lucene (Default) Result | PageRank Result |
|-----|---|---------------------------------------|
| k | http://www.foxnews.com/entertainment/2018/0 | http://www.foxnews.com/us.html |
| 1 | | nttp://www.foxnews.com/us.ntmi |
| | 2/02/stan-lee-95-feeling-great-after-return-from- | |
| | la-hospital.html | Little II |
| 2 | http://www.foxnews.com/entertainment/2018/0 | http://www.foxnews.com/science.html |
| | 2/02/stan-lee-95-feeling-great-after-return-from- | |
| | la-hospital.html | line the second second |
| 3 | http://www.foxnews.com/entertainment/2018/0 | http://www.foxnews.com/lifestyle.html |
| | 2/02/stan-lee-95-feeling-great-after-return-from- | |
| | <u>la-hospital.html</u> | |
| 4 | http://www.foxnews.com/entertainment/2018/0 | http://www.foxnews.com/food- |
| | 2/28/stan-lee-robbed-1-4m-report-says.html | <u>drink.html</u> |
| 5 | http://www.foxnews.com/tech/2017/12/20/twitt | http://www.foxnews.com/category/scien |
| | ers-purge-far-right-accounts-sparks-backlash- | ce/air-and-space.html |
| | <u>praise-and-confusion.html</u> | |
| 6 | http://www.foxnews.com/tech/2017/12/20/twitt | http://www.foxnews.com/category/healt |
| | ers-purge-far-right-accounts-sparks-backlash- | h/medical-research.html |
| | <u>praise-and-confusion.html</u> | |
| 7 | http://www.foxnews.com/us/2018/03/15/fiu- | http://www.foxnews.com/us.html |
| | bridge-collapse-sparks-reaction-from-politicians- | |
| | lawmakers.html | |
| 8 | http://www.foxnews.com/us/2018/03/15/fiu- | http://www.foxnews.com/science.html |
| | bridge-collapse-sparks-reaction-from-politicians- | |
| | lawmakers.html | |
| 9 | http://www.foxnews.com/us/2018/03/15/fiu- | http://www.foxnews.com/lifestyle.html |
| _ | bridge-collapse-sparks-reaction-from-politicians- | |
| | lawmakers.html | |
| 10 | http://www.foxnews.com/entertainment/2017/0 | http://www.foxnews.com/food- |
| | 6/17/cosby-hung-jury-judge-declares-mistrial-in- | drink.html |
| | bill-cosby-sexual-assault-case.html | |

Query 7: Austin Bombings

| Ran | Lucene (Default) Result | PageRank Result |
|-----|------------------------------------|----------------------------------|
| k | | |
| 1 | http://www.foxnews.com/us/2018/03/ | http://www.foxnews.com/tech.html |
| | 14/austin-package-bomb-victim- | |
| | remembered-as-boy-who-could-have- | |
| | accomplished-anything.html | |

| 2 | http://www.foxnews.com/us/2018/03/ | http://www.foxnews.com/category/health/cancer.h |
|----|--------------------------------------|--|
| | 14/austin-package-bomb-victim- | tml |
| | remembered-as-boy-who-could-have- | |
| | accomplished-anything.html | |
| 3 | http://www.foxnews.com/us/2018/03/ | http://www.foxnews.com/category/us/terror.html |
| | 14/austin-package-bomb-victim- | |
| | remembered-as-boy-who-could-have- | |
| | accomplished-anything.html | |
| 4 | http://www.foxnews.com/us/2018/03/ | http://www.foxnews.com/category/tech/topics/inn |
| | 15/austin-package-bombs-probably- | ovation.html |
| | pretty-simple-components-could-lead- | |
| | cops-to-suspect-experts-say.html | |
| 5 | http://www.foxnews.com/us/2018/03/ | http://www.foxnews.com/tech.html |
| | 15/austin-package-bombs-probably- | |
| | pretty-simple-components-could-lead- | |
| | cops-to-suspect-experts-say.html | |
| 6 | http://www.foxnews.com/us/2018/03/ | http://www.foxnews.com/category/health/cancer.h |
| | 15/austin-package-bombs-probably- | <u>tml</u> |
| | pretty-simple-components-could-lead- | |
| | cops-to-suspect-experts-say.html | |
| 7 | http://www.foxnews.com/us/2018/03/ | http://www.foxnews.com/category/tech/topics/inn |
| | 14/austin-package-bomb-victim- | <u>ovation.html</u> |
| | remembered-as-boy-who-could-have- | |
| | accomplished-anything.html | |
| 8 | http://www.foxnews.com/us/2018/03/ | http://www.foxnews.com/category/us/terror.html |
| | 14/austin-package-bomb-victim- | |
| | remembered-as-boy-who-could-have- | |
| | accomplished-anything.html | |
| 9 | http://www.foxnews.com/us/2018/03/ | http://www.foxnews.com/weather/2017/11/01/wi |
| | 15/austin-package-bombs-probably- | nds-from-north-korea-nuclear-test-site-may-send- |
| | pretty-simple-components-could-lead- | radiation-toward-japan/ |
| | cops-to-suspect-experts-say.html | |
| 10 | http://www.foxnews.com/us/2018/03/ | http://www.foxnews.com/category/tech/topics/sxs |
| | 15/austin-package-bombs-probably- | <u>w.html</u> |
| | pretty-simple-components-could-lead- | |
| | cops-to-suspect-experts-say.html | |

Query 8: Paris Climate Deal

| Ra | Lucene (Default) Result | PageRank Result |
|----|--|-----------------------------------|
| nk | | |
| 1 | http://www.foxnews.com/politics/2017/1 | http://www.foxnews.com/world.html |
| | 2/06/post-trump-un-climate-pact- | |
| | supporters-switch-to-state-and-local- | |
| | strategy.html | |
| 2 | http://www.foxnews.com/us/2017/09/16 | http://www.foxnews.com/tech.html |
| | /eu-officials-says-us-wont-reject-paris- | |

| | climate-treaty-but-white-house-denies- | |
|----|---|---|
| | statement.html | |
| 3 | http://www.foxnews.com/politics/2017/1 | http://www.foxnews.com/us.html |
| | 2/06/post-trump-un-climate-pact- | |
| | supporters-switch-to-state-and-local- | |
| | strategy.html | |
| 4 | http://www.foxnews.com/politics/2017/1 | http://www.foxnews.com/lifestyle.html |
| | 2/06/post-trump-un-climate-pact- | |
| | supporters-switch-to-state-and-local- | |
| | strategy.html | |
| 5 | http://www.foxnews.com/politics/2017/1 | http://www.foxnews.com/entertainment.html |
| | 2/06/post-trump-un-climate-pact- | |
| | supporters-switch-to-state-and-local- | |
| | strategy.html | |
| 6 | http://www.foxnews.com/opinion/2017/1 | http://www.foxnews.com/food-drink.html |
| | 2/12/heres-how-trump-could-make-quick- | |
| | and-complete-exit-from-paris-climate- | |
| | agreement-right-now.html | |
| 7 | http://www.foxnews.com/opinion/2017/0 | http://www.foxnews.com/auto.html |
| | 6/06/cal-thomas-secular-progressives- | |
| | reveal-again-why-are-not-fit-to-lead- | |
| | anything.html | |
| 8 | http://www.foxnews.com/world/2017/07 | http://www.foxnews.com/category/science/archa |
| | /08/g-20-summit-us-rejects-paris-climate- | eology/dinosaurs.html |
| | accord-as-19-members-reaffirm.html | |
| 9 | http://www.foxnews.com/category/us/en | http://www.foxnews.com/category/health/heart- |
| | vironment/climate-change.html | <u>health.html</u> |
| 10 | http://www.foxnews.com/world/2018/03 | http://www.foxnews.com/category/science/plane |
| | /11/french-president-pokes-at-trump-for- | <u>t-earth.html</u> |
| | <u>leaving-paris-accord.html</u> | |

Overlapping Graph For Query

| Query 1 | 0 |
|---------|---|
| Query 2 | 0 |
| Query 3 | 0 |
| Query 4 | 0 |
| Query 5 | 0 |
| Query 6 | 0 |
| Query 7 | 0 |
| Query 8 | 0 |

