Sci2 plugins:-

Note Red color signifies addition columns, additional rows or additional data

The green color signifies data in files is similar but arranged in a different order

File

Load -> Works

Read Directory Hierarchy -> Works

Split Graph to Node and Edge Files -> Works

Merge Node and Edge Files-> works

Data Preparation

Convert to Generic Publication ->

* 64 bit and 32 bit plugins generate same output files
* 64bit and 32 bit official:-
  + 64 bit has a few extra column compared to 32 bit official Associated Group, Author Identifiers, BIOSIS Citation Index, Book Author, Book Author Fullname, Book DOI, Book Group Authors, Chinese Science, Citation DB
  + 64 bit one has also an extra row.

Remove ISI duplicates

* 64 bit and 32 bit plugins generate same output files
* 64bit and 32 bit official:-
  + 64 bit has a few extra column compared to 32 bit official Associated Group, Author Identifiers, BIOSIS Citation Index, Book Author, Book Author Fullname, Book DOI, Book Group Authors, Chinese Science, Citation DB

Remove Rows with Multitudinous Fields

* 64 bit and 32 bit plugins generate same output files.
* 64bit and 32 bit official:-
  + 64 bit has a few extra column compared to 32 bit official Associated Group, Author Identifiers, BIOSIS Citation Index, Book Author, Book Author Fullname, Book DOI, Book Group Authors, Chinese Science, Citation DB
  + 64 bit one has also an extra row.

Extract Directed Networks

* 32 bit and 64 bit generate exactly the same output file.

Extract Bipartite Network

* 32 bit and 64 bit generate exactly the same output file.

Extract Paper Citation Network (green means edges and nodes upside down)

* 32 bit and 64 bit graphml files have different values for the same node.
  + 32 bit -> node 0 -> Garfield E, 1955, Science, V122, P108
  + 32 bit -> node 1 -> Wellman B, 1989, Social Structures Ne
  + 64 bit -> node 0 -> Wellman B, 1989, Social Structures Ne
  + 64 bit -> node 1 -> Garfield E, 1955, Science, V122, P108
* Paper information for 32 and 64 bit outputs have data arranged differently. I believe it is because the nodes have different values assigned to them in respective graphml files
* 64bit and 32 bit official:-
  + 32 bit official: for node id = n9, <data key="localcitationcount">1</data>
  + 64 bit: for node id = n9, <data key="localcitationcount">2</data>

Extract Author Paper Network

* 32 bit and 64 bit graphml files have different values for the same node. Same as above.
* Author Paper Information for 32 and 64 bit outputs have data arranged differently. I believe it is because the nodes have different values assigned to them in respective graphml files. Same as above.
* 64bit and 32 bit official:-
  + The output files have similar output.

Extract Co-Occurrence Network

* 32 bit and 64 bit produce the same output.

Extract Word Co-Occurrence Network- does not work for me

* The 32 and 64 bit outputs files are similar.
* 32 bit node id = n2, label = Import. Node id = n3, label = Compar.
* 32 bit edge between src =n2 and destn = n3
* 32 bit node id = n87, label = Import. Node id = n47, label = Compar.
* 32 bit edge between src =n87 and destn = n47
* The 32 bit official and 64 bit outputs files are similar.

Extract Co-Author Network

* Produces exactly the same output file for 32 bit and 64 bit output
* The 32 bit official and 64 bit outputs files are similar.

Extract Bibliographic Coupling Similarity Network

* Produces exactly the same output file for 32 bit and 64 bit output
* The 32 bit official and 64 bit outputs files are similar.

Extract Document Co-Citation Network

* The 32 bit official and 64 bit outputs files are similar.
* The nodes generated are similar.
* The edges generated are also similar but they have different edge id for the same edge.

Detect Duplicate Nodes

* 32 and 64 bit output file generated are similar
* The 32 bit official and 64 bit outputs files are similar.

Update Network by Merging Node

Preprocessing

General

* Extract Top N% Records: 32bit, 32 bit official and 64 bit plugins produce same output files.
* Extract top N rows: 32 bit, 32 bit official and 64 bit produce identical output.
* Aggregate Data: 32 bit, 32 bit official and 64 bit produce the same output files

Temporal

Slice Table by Time

* 32 bit and 64 bit produce output files with the same content but the order is not the same.
* 32 bit official and 64 bit plugin generate similar outputs and the order is also same.

Geospatial

Extract Zip Code

* 32 bit and 64 bit plugins generate similar output files.
* 32 bit official and 64 bit plugin generate similar outputs.

Topical

Reconciled Journal Names

* 32 bit and 64 bit plugins generate similar output files.
* 32 bit official and 64 bit plugin generate similar outputs.

Lowercase, Tokenize, Stem Words, Stop Words

* 32 bit and 64 bit plugins generate similar output files.
* 32 bit official and 64 bit plugin generate similar outputs.
* The path in 32 bit official output has no forwards slashes while in 64 bit has forward slashes.
* C:UsersvkarihalDesktoppresentation3rd week3-Analysis2-TopicalBurst DetectionFile to Upload

Network

Extract Top Nodes

* 32 bit and 64 bit plugin generate similar outputs
* 32 bit official and 64 bit plugin generate similar outputs.

Extract Nodes Above or below a value

* 32 bit and 64 bit plugin generate similar outputs
* 32 bit official and 64 bit plugin generate similar outputs.

Delete Isolates

* 32 bit and 64 bit plugin generate similar outputs
* 32 bit official and 64 bit plugin generate similar outputs.

Extract Top Edges

* 32 bit and 64 bit plugin generate similar outputs
* 32 bit official and 64 bit plugin generate similar outputs.

Remove Self Loops

* 32 bit and 64 bit plugin generate similar outputs
* 32 bit official and 64 bit plugin generate similar outputs.

Trim by Degree

* 32 bit and 64 bit plugin generate files which the same nodes but generate different edges.
* It can remove edges randomly. So the following makes sense.
* 32 bit has <edge id="e1" source="n2" target="n1"/> but 64 bit does not contain an edge from n2 to n1.
* 64 bit has <edge id="e0" source="n0" target="n1"/> but 32 bit does not have it.
* Same with 32 bit official and 64 bit output. The output files have the same nodes but different edges.

MST-Pathfinder Network Scaling

* 32 bit and 64 bit plugin generate similar outputs.
* 32 bit official and 64 bit plugin generate similar outputs.

Fast Pathfinder Network Scaling

* 32 bit and 64 bit plugin generate similar outputs.
* 32 bit official and 64 bit plugin generate similar outputs.

Snowball sampling (n nodes) (picks random nodes, so 32 bit and 64 bit shall be different)

* 32 bit and 64 bit plugin generate outputs files with similar skeleton.
* The nodes in two files have different values. Ex.
  + 32 bit

<node id="n0">

<data key="label">2962</data>

</node>

* + 64 bit

<node id="n0">

<data key="label">4029</data>

</node>

* The edges connect different nodes.
  + 32 bit

<edge id="e0" source="n0" target="n1"/>

* + 64 bit
* 32 bit official and 64 bit plugin generate similar skeletons but different nodes

Node Sampling (picks random nodes)

* 32 bit and 64 bit create output files with different nodes selected
* 32 bit official and 64 bit output files with different nodes selected.

Edge Sampling (picks random edges)

* 32 bit and 64 bit create output files with random nodes and edges.
* 32 bit official and 64 bit create output files with random nodes and edges..

Dichotomize

Multipartite joining

Merge two Networks

Analysis

Temporal (Burst Detection)

* The out files have similar data. The analysis files for 32 bit and 64 bit have data arranged differently.

Geospatial

Generic Geocoder

* 32 bit and 64 bit plugin generate similar outputs

Bing Geocoder

Topical (Burst Detection)

* The out files have similar data. The analysis files for 32 bit and 64 bit have data arranged differently.
* The decimal precision for weight variable for 32 bit and 64 bit are different.

Network

NAT

* Both 32 bit and 64 bit plugin generate the same output.

Un-weighted & Undirected

Node Degree

* + Both 32 bit and 64 bit plugin generate the same output.

K – Nearest Neighbor

* + Both 32 bit and 64 bit plugin generate the same output.

Node Betweenness Centrality

* + Both 32 bit and 64 bit plugin generate the same output. But the edge id that connect he same pair of nodes are different.
  + 32 bit
    - <edge id="e1" source="n4" target="n3" />
  + 64 bit
    - <edge id="e3" source="n4" target="n3" />

Extract K core

* + Both 32 bit official Sci2 and 64 bit plugin generate the same output.

Annotate K-Coreness

* + Both 32 bit official Sci2 and 64 bit plugin generate the same output.

Blondel Community Detection

* Works with the official 32 bit Sci2. But throws errors with the new 32bit and 64 bit build.
* Error: org.cishell.utilities.AlgorithmNotFoundException: Unable to find an algorithm that satisfied the following filter:

(service.pid=edu.iu.nwb.shared.blondelexecutable)

Louvain Community Detection (with resolution parameter)

* + Both 32 and 64 bit generate same output files.

Louvain Multilevel Refinement Community Detection

* + Both 32 and 64 bit generate same output files.

SLM Community Detection

* + Both 32 and 64 bit generate same output files.

HITS

* + Both 32 and 64 bit generate same output files.

Weighted & Undirected

Node Betweenness Centrality

* + Both 32 bit official Sci2 and 64 bit plugin generate the same output. But the edge id that connect he same pair of nodes are different.
  + 32 bit
    - <edge id="e1" source="n4" target="n3" />
  + 64 bit
    - <edge id="e3" source="n4" target="n3" />

Blondel Community Detection

* Works with the 32 bit official Sci2. But throws errors with the new 32bit and 64 bit build.
* Error: org.cishell.utilities.AlgorithmNotFoundException: Unable to find an algorithm that satisfied the following filter:

(service.pid=edu.iu.nwb.shared.blondelexecutable)

Louvain Community Detection (with resolution parameter)

* + Both 32 bit official Sci2 and 64 bit plugin generate the same output.

Louvain Multilevel Refinement Community Detection

* + Both 32 bit official Sci2 and 64 bit plugin generate the same output.

SLM Community Detection

* + Both 32 bit official Sci2 and 64 bit plugin generate the same output.

HITS

* + Both 32 bit official Sci2 and 64 bit plugin generate the same output.

Unweighted & Directed

Node Indegree

* + Both 32 bit official and 64 bit generate same output files.

Node Out Degree

* + Both 32 bit official and 64 bit generate same output files.

K NN

* + Both 32 bit official and 64 bit generate same output files.

Node Betweenness Centrality

* + Both official 32 bit and 64 bit plugin generate the same output.
  + 32 bit <edge id="e1" source="n4" target="n3" />
  + 64 bit <edge id="e3" source="n13" target="n2" />

Weak Component Cluster

* + Official 32 bit and 64 bit plugin generate file with similar nodes but the edges connect different nodes.
  + 32 bit -> edge id e1 and 64 bit-> edge id e15
  + 32 bit -> edge id e0 and 64 bit-> edge id e18

Strong Component Cluster

* + Both Official 32 and 64 bit generate same output files.

Blondel Community Detection

* + org.cishell.utilities.AlgorithmNotFoundException: Unable to find an algorithm that satisfied the following filter:

(service.pid=edu.iu.nwb.shared.blondelexecutable)

Louvain Community Detection (with resolution parameter)

* + Both official 32 and 64 bit generate same output files.

Louvain Multilevel Refinement Community Detection

* + Both official 32 and 64 bit generate same output files.

SLM Community Detection

* + Both official 32 and 64 bit generate same output files.

Extract K-Core

* + Both Official 32 bit and 64 bit generate same output files.

Annotate K-Coreness

* + Both Official 32 and 64 bit generate same output files.

HITS

* + Both 32 bit official and 64 bit generate same output file

Page Rank

* + Both 32 bit official and 64 bit generate same output files.

Weighted & Directed

Blondel Community Detection

* Works with the official Sci2. But throws errors with the new 32bit and 64 bit build.
* Error: org.cishell.utilities.AlgorithmNotFoundException: Unable to find an algorithm that satisfied the following filter:

(service.pid=edu.iu.nwb.shared.blondelexecutable)

Louvain Community Detection (with resolution parameter)

* + Both 32 bit official Sci2 and 64 bit plugin generate the same output.

Louvain Multilevel Refinement Community Detection

* + Both 32 bit official Sci2 and 64 bit plugin generate the same output.

SLM Community Detection

* + Both 32 bit official Sci2 and 64 bit plugin generate the same output.

HITS

* + Both 32 bit official Sci2 and 64 bit plugin generate the same output.

PageRank

* + Both 32 bit official Sci2 and 64 bit plugin generate the same output.

Modeling

Networks

TARL

* + Both 32 bit official Sci2 and 64 bit plugin generate the same output.

Visualization

Temporal (Temporal Bar Graph)

* 32 bit and 64 bit plugin generate identical files
* The visualization created by 32 bit and 64 bit plugins are identical

Geospatial

Proportional Symbol Map

* 32 bit and 64 bit create similar visualization.

Choropleth Map

* 32 bit and 64 bit create similar visualization.

Geospatial Network Layout with Base Map

Topical

Map of Science via Journals

Map of Science via 554 Fields

Networks

Radial Tree Graph

* Generate similar visualization.

Radial Tree Graph with Annotation

* Generate similar visualization.

Tree View

* Generate similar visualization.

Tree Map

* Generate similar visualization.

Fruchterman-Reingold with Annotation

* Generate similar visualization.

Drl

* 32 bit and 64 bit plugin generate similar output files.

Bipartite Network Graph

* 32 bit and 64 bit plugin generate similar output files.