

ScaleGraph/Python Graph Algorithm API

ScaleGraph/Pythonが提供するGraph Algorithm APIについて記載します。

[ScaleGraph/Python Graph Algorithm API](#)

[Using API](#)

[PageRank](#)

[Function Definition](#)

[Extra Options](#)

[Degree Distribution](#)

[Function Definition](#)

[Extra Options](#)

[Betweenness Centrality](#)

[Function Definition](#)

[Extra Options](#)

[HyperANF](#)

[Function Definition](#)

[Extra Options](#)

[Strongly Connected Component](#)

[Function Definition](#)

[Extra Options](#)

[Minimum Spanning Tree](#)

[Function Definition](#)

[Extra Options](#)

[Max Flow](#)

[Function Definition](#)

[Extra Options](#)

[Spectral Clustering](#)

[Function Definition](#)

[Extra Options](#)

Using API

下記のようにscalegraphモジュールのimportで使えます。

```
$ PYTHONPATH=/path/to/project/of/scalegraph-dev/src/python python3  
import scalegraph
```

PageRank

Function Definition

```
scalegraph.pagerank(input=None, input_path=None, input_fs=scalegraph.OS,  
input_rmat_scale=8, output_path=None, output_fs=scalegraph.OS, extra_options=[])
```

- input: scalegraph.FILEまたはscalegraph.RMATの値をとります。scalegraph.FILEの指定した場合、入力グラフデータとしてファイルシステム（UnixファイルシステムまたはHDFS）からデータを読み込みます。scalegraph.RMATを指定した場合、入力グラフデータは内部で自動生成します。
- input_path: inputがscalegraph.FILEのときに利用するファイルのパスを指定します。
- input_fs: inputがscalegraph.FILEのときに利用するファイルシステムの種別を指定します。scalegraph.OSであればUnixファイルシステム、scalegraph.HDFSであればHDFSとなります。
- input_rmat_scale: inputがscalegraph.RMATのときのデータ生成パラメータです。
- output_path: 結果の出力先となるファイルパスを指定します。
- output_fs: 結果の出力先のファイルシステムの種別を指定します。scalegraph.OSであればUnixファイルシステム、scalegraph.HDFSであればHDFSとなります。
- extra_options: 下記のアルゴリズム固有オプションを記述します。

Extra Options

下記、計算パラメータを記述することができます。記述しない場合はデフォルト値が利用されます。

- --pr-damping=0.95
- --pr-eps=0.002
- --pr-niter=50

Degree Distribution

Function Definition

`scalegraph.degreedistribution(input=None, input_path=None, input_fs=scalegraph.OS, input_rmat_scale=8, output_path=None, output_fs=scalegraph.OS, extra_options=[])`

- `input`: `scalegraph.FILE`または`scalegraph.RMAT`の値をとります。 `scalegraph.FILE`の指定した場合、入力グラフデータとしてファイルシステム（UnixファイルシステムまたはHDFS）からデータを読み込みます。 `scalegraph.RMAT`を指定した場合、入力グラフデータは内部で自動生成します。
- `input_path`: `input`が`scalegraph.FILE`のときに利用するファイルのパスを指定します。
- `input_fs`: `input`が`scalegraph.FILE`のときに利用するファイルシステムの種別を指定します。 `scalegraph.OS`であればUnixファイルシステム、 `scalegraph.HDFS`であればHDFSとなります。
- `input_rmat_scale`: `input`が`scalegraph.RMAT`のときのデータ生成パラメータです。
- `output_path`: 結果の出力先となるファイルパスを指定します。
- `output_fs`: 結果の出力先のファイルシステムの種別を指定します。 `scalegraph.OS`であればUnixファイルシステム、 `scalegraph.HDFS`であればHDFSとなります。
- `extra_options`: 下記のアルゴリズム固有オプションを記述します。

Extra Options

下記、計算パラメータを記述することができます。記述しない場合はデフォルト値が利用されます。

- `--output-data-file-header-weight=not-read`

Betweenness Centrality

Function Definition

`scalegraph.pagerank(input=None, input_path=None, input_fs=scalegraph.OS, input_rmat_scale=8, output_path=None, output_fs=scalegraph.OS, extra_options=[])`

- `input`: `scalegraph.FILE`または`scalegraph.RMAT`の値をとります。 `scalegraph.FILE`の指

定した場合、入力グラフデータとしてファイルシステム（UnixファイルシステムまたはHDFS）からデータを読み込みます。scalegraph.RMATを指定した場合、入力グラフデータは内部で自動生成します。

- input_path: inputがscalegraph.FILEのときに利用するファイルのパスを指定します。
- input_fs: inputがscalegraph.FILEのときに利用するファイルシステムの種別を指定します。scalegraph.OSであればUnixファイルシステム、scalegraph.HDFSであればHDFSとなります。
- input_rmat_scale: inputがscalegraph.RMATのときのデータ生成パラメータです。
- output_path: 結果の出力先となるファイルパスを指定します。
- output_fs: 結果の出力先のファイルシステムの種別を指定します。scalegraph.OSであればUnixファイルシステム、scalegraph.HDFSであればHDFSとなります。
- extra_options: 下記のアルゴリズム固有オプションを記述します。

Extra Options

下記、計算パラメータを記述することができます。記述しない場合はデフォルト値が利用されます。

- --bc-weighted
- --bc-directed
- --bc-source
- --bc-delta
- --bc-normalize
- --bc-linearScale
- --bc-exactBC

HyperANF

Function Definition

```
scalegraph.pagerank(input=None, input_path=None, input_fs=scalegraph.OS,  
input_rmat_scale=8, output_path=None, output_fs=scalegraph.OS, extra_options=[])
```

- input: scalegraph.FILEまたはscalegraph.RMATの値をとります。scalegraph.FILEの指定した場合、入力グラフデータとしてファイルシステム（UnixファイルシステムまたはHDFS）からデータを読み込みます。scalegraph.RMATを指定した場合、入力グラフデータは内部で自動生成します。
- input_path: inputがscalegraph.FILEのときに利用するファイルのパスを指定しま

す。

- `input_fs`: `input`が`scalegraph.FILE`のときに利用するファイルシステムの種別を指定します。`scalegraph.OS`であればUnixファイルシステム、`scalegraph.HDFS`であればHDFSとなります。
- `input_rmat_scale`: `input`が`scalegraph.RMAT`のときのデータ生成パラメータです。
- `output_path`: 結果の出力先となるファイルパスを指定します。
- `output_fs`: 結果の出力先のファイルシステムの種別を指定します。`scalegraph.OS`であればUnixファイルシステム、`scalegraph.HDFS`であればHDFSとなります。
- `extra_options`: 下記のアルゴリズム固有オプションを記述します。

Extra Options

下記、計算パラメータを記述することができます。記述しない場合はデフォルト値が利用されます。

- `--hanf-niter`
- `--hanf-b`

Strongly Connected Component

Function Definition

```
scalegraph.pagerank(input=None, input_path=None, input_fs=scalegraph.OS,  
input_rmat_scale=8, output_path=None, output_fs=scalegraph.OS, extra_options=[])
```

- `input`: `scalegraph.FILE`または`scalegraph.RMAT`の値をとります。`scalegraph.FILE`の指定した場合、入力グラフデータとしてファイルシステム（UnixファイルシステムまたはHDFS）からデータを読み込みます。`scalegraph.RMAT`を指定した場合、入力グラフデータは内部で自動生成します。
- `input_path`: `input`が`scalegraph.FILE`のときに利用するファイルのパスを指定します。
- `input_fs`: `input`が`scalegraph.FILE`のときに利用するファイルシステムの種別を指定します。`scalegraph.OS`であればUnixファイルシステム、`scalegraph.HDFS`であればHDFSとなります。
- `input_rmat_scale`: `input`が`scalegraph.RMAT`のときのデータ生成パラメータです。
- `output_path`: 結果の出力先となるファイルパスを指定します。
- `output_fs`: 結果の出力先のファイルシステムの種別を指定します。`scalegraph.OS`であればUnixファイルシステム、`scalegraph.HDFS`であればHDFSとなります。
- `extra_options`: 下記のアルゴリズム固有オプションを記述します。

Extra Options

下記、計算パラメータを記述することができます。記述しない場合はデフォルト値が利用されます。

- --scc-directed
- --scc-niter

Minimum Spanning Tree

Function Definition

```
scalegraph.pagerank(input=None, input_path=None, input_fs=scalegraph.OS,  
input_rmat_scale=8, output_path=None, output_fs=scalegraph.OS, extra_options=[])
```

- input: scalegraph.FILEまたはscalegraph.RMATの値をとります。scalegraph.FILEの指定した場合、入力グラフデータとしてファイルシステム（UnixファイルシステムまたはHDFS）からデータを読み込みます。scalegraph.RMATを指定した場合、入力グラフデータは内部で自動生成します。
- input_path: inputがscalegraph.FILEのときに使用するファイルのパスを指定します。
- input_fs: inputがscalegraph.FILEのときに使用するファイルシステムの種別を指定します。scalegraph.OSであればUnixファイルシステム、scalegraph.HDFSであればHDFSとなります。
- input_rmat_scale: inputがscalegraph.RMATのときのデータ生成パラメータです。
- output_path: 結果の出力先となるファイルパスを指定します。
- output_fs: 結果の出力先のファイルシステムの種別を指定します。scalegraph.OSであればUnixファイルシステム、scalegraph.HDFSであればHDFSとなります。
- extra_options: 下記のアルゴリズム固有オプションを記述します。

Extra Options

TBD // オリジナルのアルゴリズムが正常に動かないため

Max Flow

Function Definition

`scalegraph.pagerank(input=None, input_path=None, input_fs=scalegraph.OS, input_rmat_scale=8, output_path=None, output_fs=scalegraph.OS, extra_options=[])`

- `input`: `scalegraph.FILE`または`scalegraph.RMAT`の値をとります。 `scalegraph.FILE`の指定した場合、入力グラフデータとしてファイルシステム（UnixファイルシステムまたはHDFS）からデータを読み込みます。 `scalegraph.RMAT`を指定した場合、入力グラフデータは内部で自動生成します。
- `input_path`: `input`が`scalegraph.FILE`のときに使用するファイルのパスを指定します。
- `input_fs`: `input`が`scalegraph.FILE`のときに使用するファイルシステムの種別を指定します。 `scalegraph.OS`であればUnixファイルシステム、 `scalegraph.HDFS`であればHDFSとなります。
- `input_rmat_scale`: `input`が`scalegraph.RMAT`のときのデータ生成パラメータです。
- `output_path`: 結果の出力先となるファイルパスを指定します。
- `output_fs`: 結果の出力先のファイルシステムの種別を指定します。 `scalegraph.OS`であればUnixファイルシステム、 `scalegraph.HDFS`であればHDFSとなります。
- `extra_options`: 下記のアルゴリズム固有オプションを記述します。

Extra Options

下記、計算パラメータを記述することができます。記述しない場合はデフォルト値が利用されます。

- `--mf-source-id`
- `--mf-sink-id`
- `--mf-eps`
- `--mf-recursionLimit`

Spectral Clustering

Function Definition

`scalegraph.pagerank(input=None, input_path=None, input_fs=scalegraph.OS, input_rmat_scale=8, output_path=None, output_fs=scalegraph.OS, extra_options=[])`

- `input`: `scalegraph.FILE`または`scalegraph.RMAT`の値をとります。 `scalegraph.FILE`の指定した場合、入力グラフデータとしてファイルシステム（Unixファイルシステム

またはHDFS) からデータを読み込みます。scalegraph.RMATを指定した場合、入力グラフデータは内部で自動生成します。

- input_path: inputがscalegraph.FILEのときに使用するファイルのパスを指定します。
- input_fs: inputがscalegraph.FILEのときに使用するファイルシステムの種別を指定します。scalegraph.OSであればUnixファイルシステム、scalegraph.HDFSであればHDFSとなります。
- input_rmat_scale: inputがscalegraph.RMATのときのデータ生成パラメータです。
- output_path: 結果の出力先となるファイルパスを指定します。
- output_fs: 結果の出力先のファイルシステムの種別を指定します。scalegraph.OSであればUnixファイルシステム、scalegraph.HDFSであればHDFSとなります。
- extra_options: 下記のアルゴリズム固有オプションを記述します。

Extra Options

下記、計算パラメータを記述することができます。記述しない場合はデフォルト値が利用されます。

TBD // オリジナルのアルゴリズムが正常に動かないため