

Package ‘RWgraph’

September 21, 2023

Title Random Walks on Graphs Representing a Transactional Network

Version 1.0.0

Description Random walk functions to extract new variables based on clients transactional behaviour. For more details, see Ed-
din et al. (2021) <[arXiv:2112.07508v3](#)> and Oliveira et al. (2021) <[arXiv:2102.05373v2](#)>.

License GPL (>= 2)

Encoding UTF-8

RoxygenNote 7.2.3

Imports igraph

Depends R (>= 2.10)

LazyData true

NeedsCompilation no

Author Mafalda Sá Ferreira [aut, cre],
Regina Bispo [ctb],
FCT, I.P. [fnd] (under the scope of the projects UIDB/00297/2020 and
UIDP/00297/2020 (NovaMath))

Maintainer Mafalda Sá Ferreira <msm.ferreira@campus.fct.unl.pt>

Repository CRAN

Date/Publication 2023-09-21 18:30:03 UTC

R topics documented:

clients_small_example	2
info_client	2
mean_rw_client	3
profiles	4
rw_client	5
transactions	6
transactions_small_example	6

Index	7
--------------	----------

clients_small_example *Clients' information for a small example*

Description

A dataset containing information about 20 clients of a certain bank.

Usage

```
data(clients_small_example)
```

Format

A data frame with 20 rows and 9 variables

Details

- age, numeric. Age of the client in years.
- antiquity_age, numeric. Age of the account in years.
- gender, boolean. Gender of the client.
- occupation, numeric. Occupation of the client.
- nationality, character. Country of birth of the client (labelled in ISO-CODE 2).
- residence, character. Country of residence of the client (labelled in ISO-CODE 2).
- pep_flag, boolean. Indicator whether the client is involved in political activities (1) or not (0).
- sar_flag, boolean. Indicator whether the client was involved in a reported transaction (1) or not (0).
- customer_id, character. ID of the client's account.

info_client *Random walk metrics for each client*

Description

Computes the metrics of the generated random walks for every client in the dataframe using the function 'mean_rw_client'.

Usage

```
info_client(g, data)
```

Arguments

<code>g</code>	The input graph. Transactional graph containing the amount (in monetary unit) as the attribute of each edge. The vertices must be the clients IDs.
<code>data</code>	Dataframe with information of the clients. It should include a column with the clients IDs named "customer_id" and the alert label named "sar_flag" that must be a boolean variable.

Value

A dataframe with the clients IDs and the computed metrics (minimum, mean and maximum for both the number of steps and total transacted amount) for the random walks starting in each client.

References

Eddin, A. N., Bono, J., Aparício, D., Polido, D., Ascensao, J. T., Bizarro, P., and Ribeiro, P. (2021). Anti-money laundering alert optimization using machine learning with graphs. arXiv preprint arXiv:2112.07508.

Examples

```
g <- igraph::graph_from_data_frame(d = transactions_small_example[, 1:3], directed = TRUE)
info_client(g, data = clients_small_example)
```

mean_rw_client	<i>Metrics for multiple random walks</i>
----------------	--

Description

Computes metrics for 50 generated random walks using the function 'rw_client'.

Usage

```
mean_rw_client(v, g, data)
```

Arguments

<code>v</code>	The initial vertex of the input graph.
<code>g</code>	The input graph. It should be a transactional graph with the amount as the attribute of each edge. The vertices must be the clients IDs.
<code>data</code>	Dataframe with information of the clients. It should include a column with the clients IDs named "customer_id" and the alert label named "sar_flag" that must be a boolean variable.

Value

A vector with the minimum, mean and maximum for both the number of steps and total transacted amount in the random walks calculated.

References

Eddin, A. N., Bono, J., Aparício, D., Polido, D., Ascensao, J. T., Bizarro, P., and Ribeiro, P. (2021). Anti-money laundering alert optimization using machine learning with graphs. arXiv preprint arXiv:2112.07508.

Examples

```
g <- igraph::graph_from_data_frame(d = transactions_small_example[, 1:3], directed = TRUE)
v <- transactions_small_example[1, 1]
mean_rw_client(v, g, data = clients_small_example)
```

profiles

Clients' information

Description

A dataset containing information about 3973 clients of a certain bank.

Usage

```
data(profiles)
```

Format

A data frame with 3973 rows and 9 variables

Details

- `age`, numeric. Age of the client in years.
- `antiquity_age`, numeric. Age of the account in years.
- `gender`, boolean. Gender of the client.
- `occupation`, numeric. Occupation of the client.
- `nationality`, character. Country of birth of the client (labelled in ISO-CODE 2).
- `residence`, character. Country of residence of the client (labelled in ISO-CODE 2).
- `pep_flag`, boolean. Indicator whether the client is involved in political activities (1) or not (0).
- `sar_flag`, boolean. Indicator whether the client was involved in a reported transaction (1) or not (0).
- `customer_id`, character. ID of the client's account.

rw_client	<i>Random walk simulation</i>
-----------	-------------------------------

Description

Computes a random walk path for a given client.

Usage

```
rw_client(v, g, data)
```

Arguments

v	The initial vertex of the input graph.
g	The input graph. It should be a transactional graph with the amount as the attribute of each edge. The vertices must be the clients IDs.
data	Dataframe with information of the clients. It should include a column with the clients IDs named "customer_id" and the alert label named "sar_flag" that must be a boolean variable.

Value

A vector with the number of steps taken in the random walk and the total transacted amount in it.

References

Eddin, A. N., Bono, J., Aparício, D., Polido, D., Ascensao, J. T., Bizarro, P., and Ribeiro, P. (2021). Anti-money laundering alert optimization using machine learning with graphs. arXiv preprint arXiv:2112.07508.

Examples

```
g <- igraph::graph_from_data_frame(d = transactions_small_example[, 1:3], directed = TRUE)
v <- transactions_small_example[1, 1]
rw_client(v, g, data = clients_small_example)
```

transactions

Transactions' information

Description

A dataset containing information about 15379 transactions of a certain bank.

Usage

```
data(transactions)
```

Format

A data frame with 15379 rows and 5 variables

Details

- nameOrig, character. ID of the client that initiated the transaction.
 - nameDest, character. ID of the client that received the transaction.
 - amount, numeric. Amount of money involved in the transaction in euros (€).
 - isFraud, boolean. Indicator whether the transaction was reported (1) or not (0).
 - transactionDate, character. Date of the transaction.
-

transactions_small_example

Transactions' information for a small example

Description

A dataset containing information about 10 transactions of a certain bank.

Usage

```
data(transactions_small_example)
```

Format

A data frame with 10 rows and 5 variables

Details

- nameOrig, character. ID of the client that initiated the transaction.
- nameDest, character. ID of the client that received the transaction.
- amount, numeric. Amount of money involved in the transaction in euros (€).
- isFraud, boolean. Indicator whether the transaction was reported (1) or not (0).
- transactionDate, character. Date of the transaction.

Index

clients_small_example, 2

info_client, 2

mean_rw_client, 3

profiles, 4

rw_client, 5

transactions, 6

transactions_small_example, 6