`triversity`: an `R` package to compute diversity measures on multipartite graphs

Robin Lamarche-Perrin

Centre national de la recherche scientifique
Institut des systèmes complexes de Paris Île-de-France
Laboratoire d'informatique de Paris 6

Meeting of the task-force on diversity measures
for the ANR AIgoDiv Project
23rd of October, 2017, in Paris

Install and load `triversity`

- `triversity` is an R package for the computation of diversity measures on tripartite graphs.
- It implements the parametrized family of "true diversity" measures, notably containing the richness, the Shannon entropy, the Herfindahl-Hirschman index, and the Berger-Parker index.
- It applies these measures on probability distributions resulting from random walks between the levels of tripartite graphs

Published on CRAN:
`https://cran.r-project.org/web/packages/triversity`

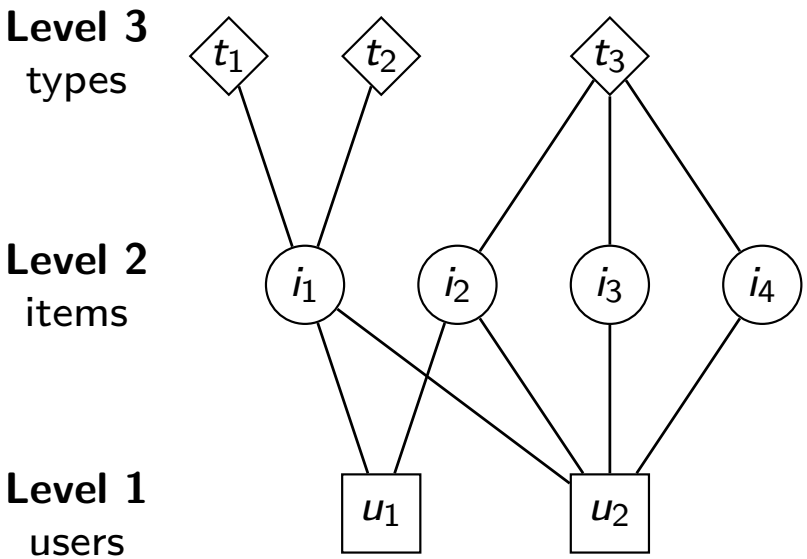
Source on GitHub:
`https://github.com/Lamarche-Perrin/triversity`

To install and load:

```
install.packages ('triversity')
```

```
library ('triversity')
```

Load a tripartite graph

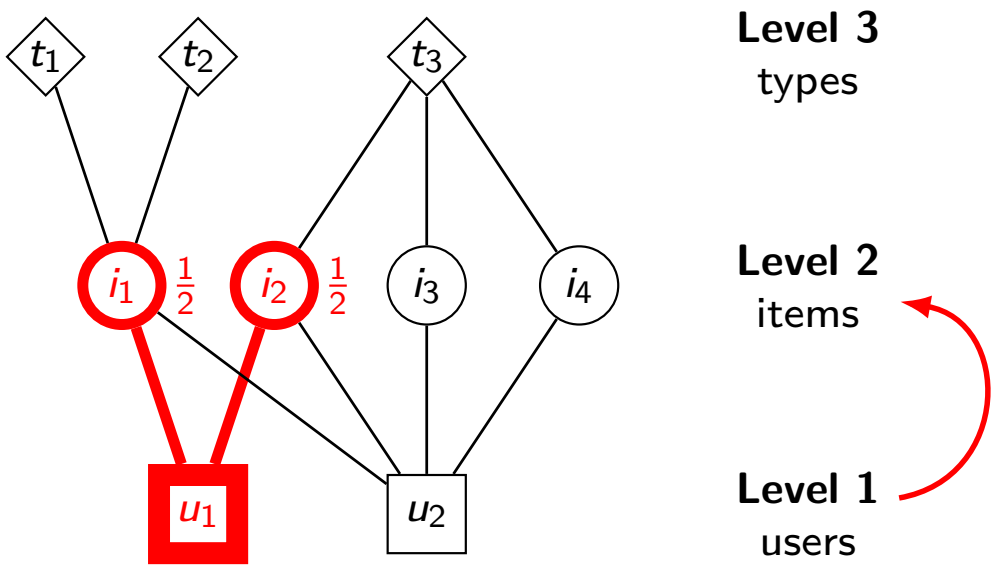



```
read.table ('tripartite_example.csv')
```

```
##    V1 V2 V3 V4
## 1   1 u1  2 i1
## 2   1 u1  2 i2
## 3   1 u2  2 i1
## 4   1 u2  2 i2
## 5   2 i3  1 u2
## 6   2 i4  1 u2
## 7   2 i1  3 t1
## 8   2 i1  3 t2
## 9   2 i2  3 t3
## 10  3 t3  2 i3
## 11  3 t3  2 i4
```

```
example <- get_multipartite ('tripartite_example.csv')
```

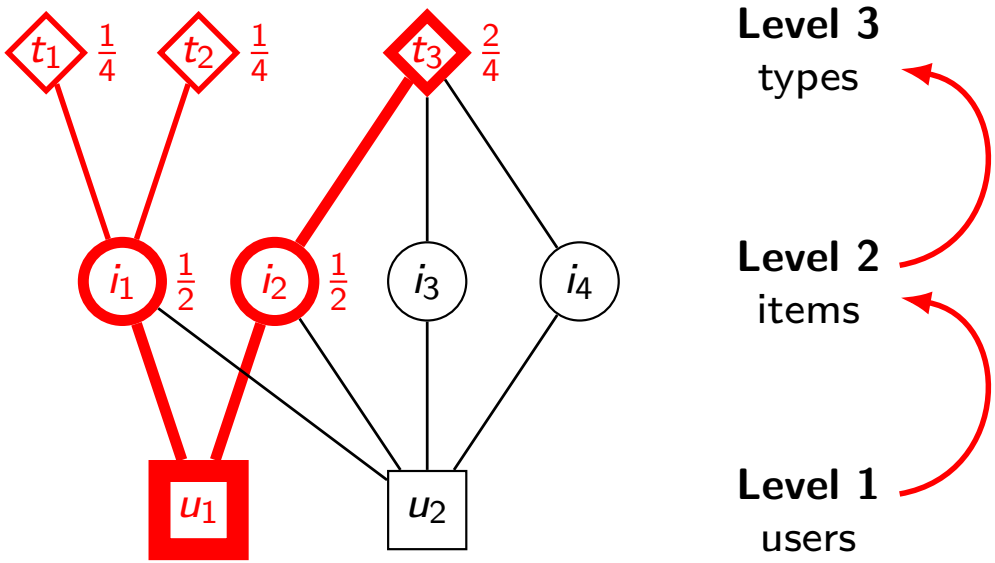
Item-diversity of a given user




```
get_diversity_from_path (
    graph = example, path = c(1,2),
    initial_node = 'u1',
    measure = c('richness', 'entropy', 'herfindahl')
)
```

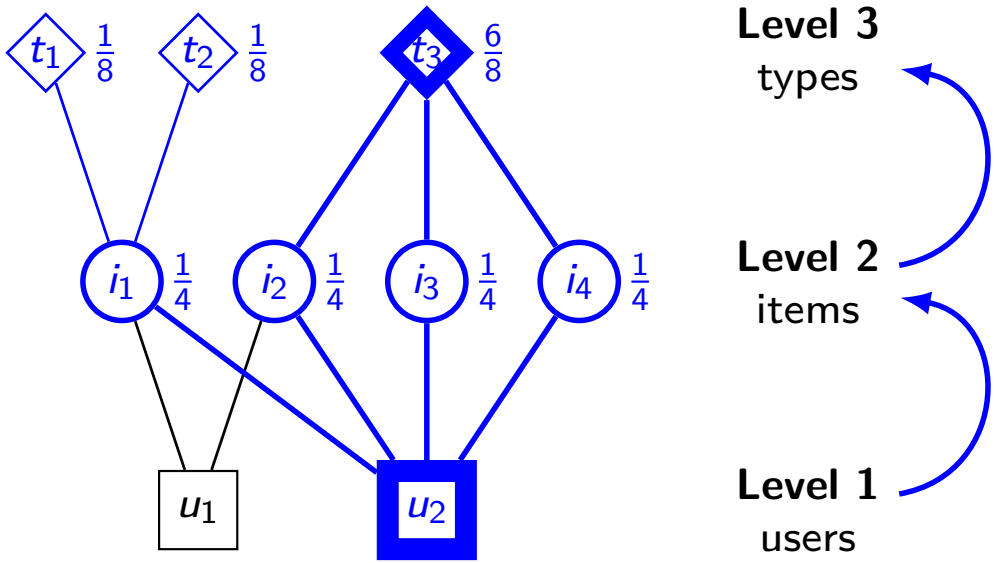
```
##   richness    entropy herfindahl
##        2.0        1.0        0.5
```

Type-diversity of a given user




```
get_diversity_from_path (
    graph = example, path = c(1,2,3),
    initial_node = 'u1',
    measure = c('richness', 'entropy', 'herfindahl')
)

##   richness    entropy herfindahl
##      3.000      1.500      0.375
```

Type-diversity of all users

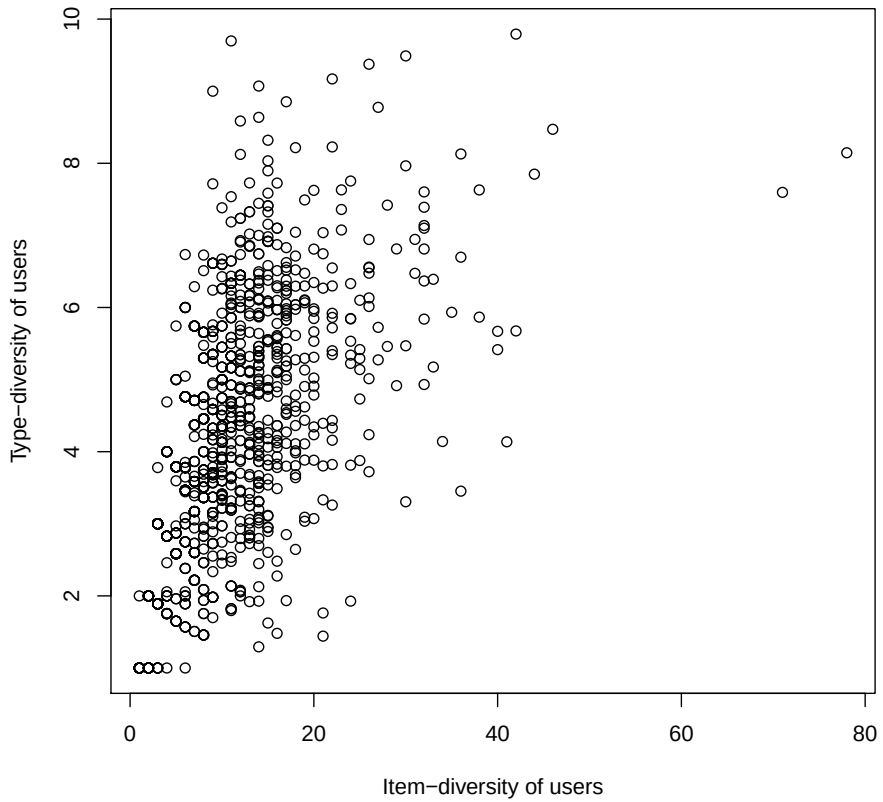



```
get_diversity_from_path (
    graph = example, path = c(1,2,3),
    measure = c('richness', 'entropy', 'herfindahl')
)
```

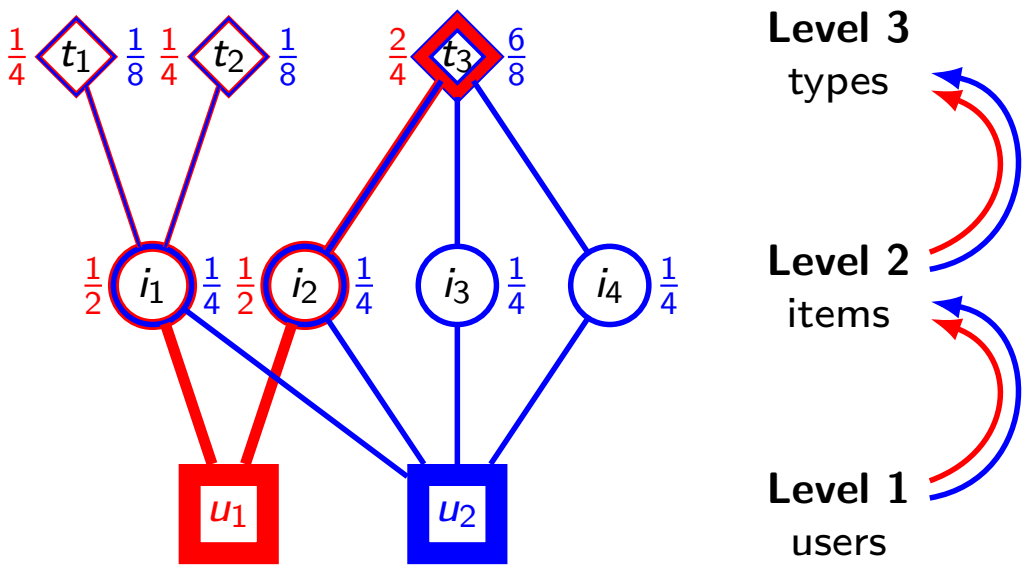
```
##    richness  entropy herfindahl
## u1        3 1.500000    0.37500
## u2        3 1.061278    0.59375
```

Item-diversity vs. Type-diversity of all users

```
automotive <- get_multipartite ('tripartite_automotive_sample.csv')
user_item_div <- get_diversity_from_path (
    graph=automotive, path=c(1,2), order=1)
user_type_div <- get_diversity_from_path (
    graph=automotive, path=c(1,2,3), order=1)
plot (user_item_div, user_type_div,
      xlab='Item-diversity of users', ylab='Type-diversity of users')
```



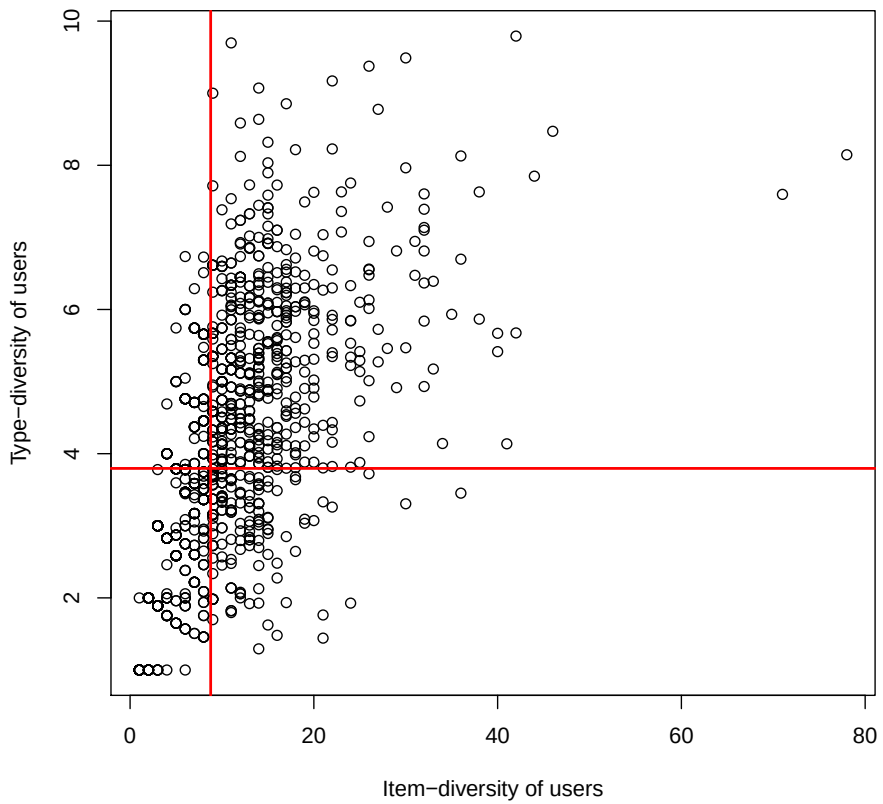
Mean of individual type-diversities of users



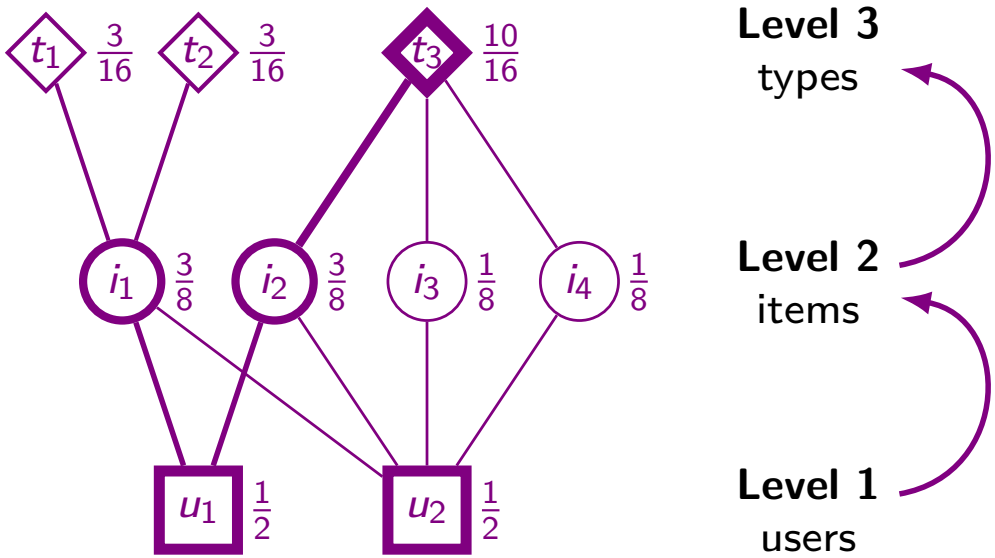

```
get_diversity_from_path (
    graph = example, path = c(1,2,3),
    type = 'mean',
    measure = c('richness', 'entropy', 'herfindahl')
)

##   richness    entropy herfindahl
##  3.0000000  1.2806391  0.4718647
```

Item-diversity vs. Type-diversity of all users

```
ind_user_item_div <- get_diversity_from_path (
    graph=automotive, path=c(1,2), type='mean', order=1)
ind_user_type_div <- get_diversity_from_path (
    graph=automotive, path=c(1,2,3), type='mean', order=1)
plot (user_item_div, user_type_div,
    xlab='Item-diversity of users', ylab='Type-diversity of users')
abline (v=ind_user_item_div, h=ind_user_type_div, col='red', lwd=2)
```



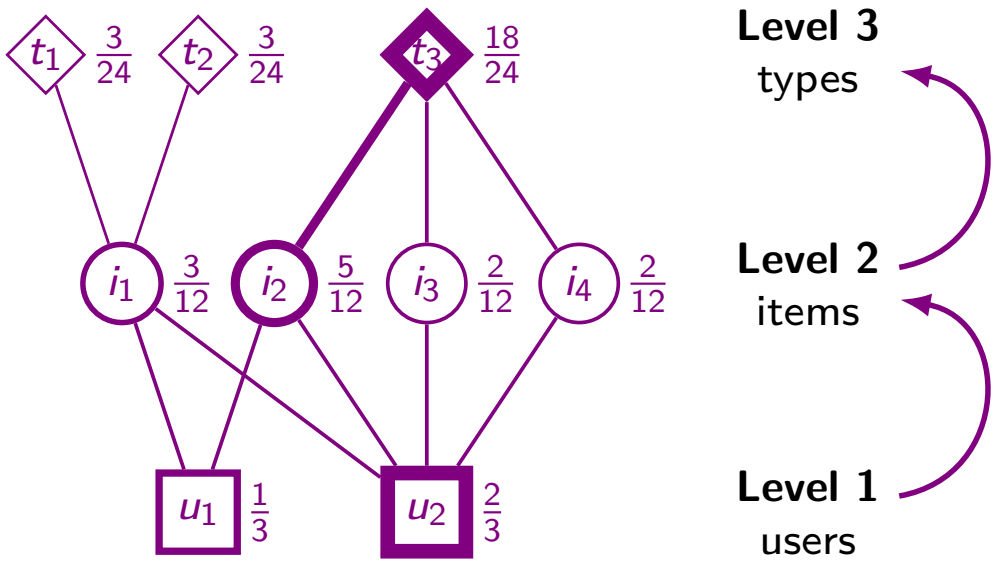
Collective type-diversity of users (uniform initial distribution)




```
get_diversity_from_path (
    graph = example, path = c(1,2,3),
    type = 'collective',
    measure = c('richness', 'entropy', 'herfindahl')
)
```

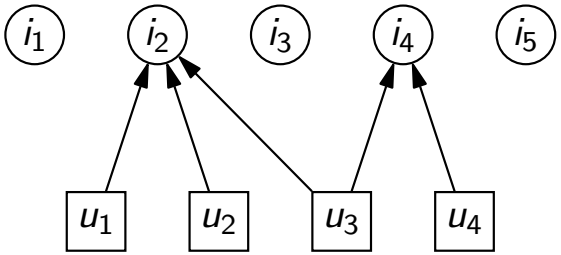
```
##   richness    entropy herfindahl
##  3.0000000  1.3294340  0.4609375
```

Collective type-diversity of users (biased initial distribution)

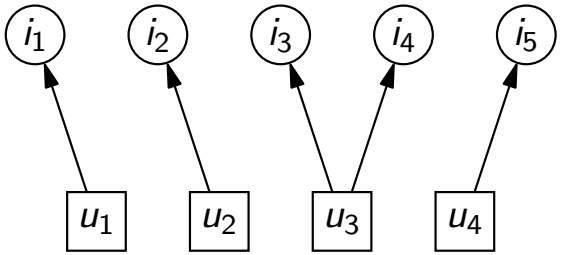



```
get_diversity_from_path (
    graph = example, path = c(1,2,3),
    type = 'collective',
    initial_distribution = c(1/3, 2/3),
    measure = c('richness', 'entropy', 'herfindahl')
)

##   richness    entropy herfindahl
##   3.000000   1.251629   0.500000
```

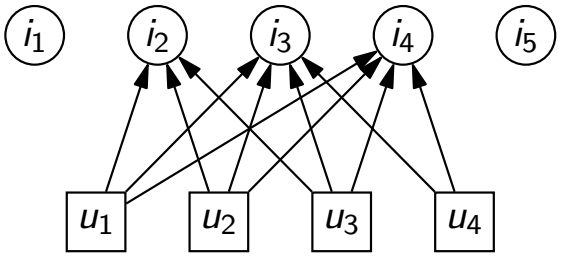
Individual diversity vs. Collective diversity



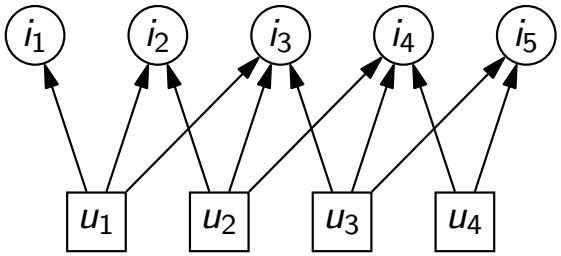
Weak individual diversity
Weak collective diversity



Weak individual diversity
Strong collective diversity



Strong individual diversity
Weak collective diversity



Strong individual diversity
Strong collective diversity

Individual vs. Collective and Item vs. Type

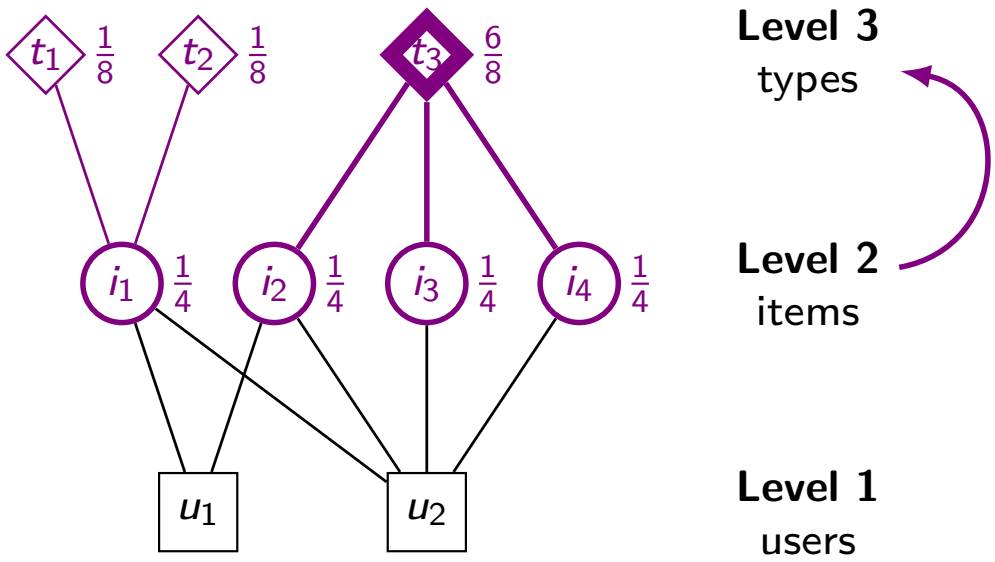
```
col_user_item_div <- get_diversity_from_path (
    graph=automotive, path=c(1,2), type='collective', order=1)
col_user_type_div <- get_diversity_from_path (
    graph=automotive, path=c(1,2,3), type='collective', order=1)
```

	Mean of individual ...	Collective ...
... item-diversity of users	8.7701001	3537.7344181
... type-diversity of users	3.7958606	10.6452442

```
ind_item_rich <- get_diversity_from_path (
    graph=automotive, path=c(1,2), type='mean', order=0)
ind_type_rich <- get_diversity_from_path (
    graph=automotive, path=c(1,2,3), type='mean', order=0)
col_item_rich <- get_diversity_from_path (
    graph=automotive, path=c(1,2), type='collective', order=0)
col_type_rich <- get_diversity_from_path (
    graph=automotive, path=c(1,2,3), type='collective', order=0)
```

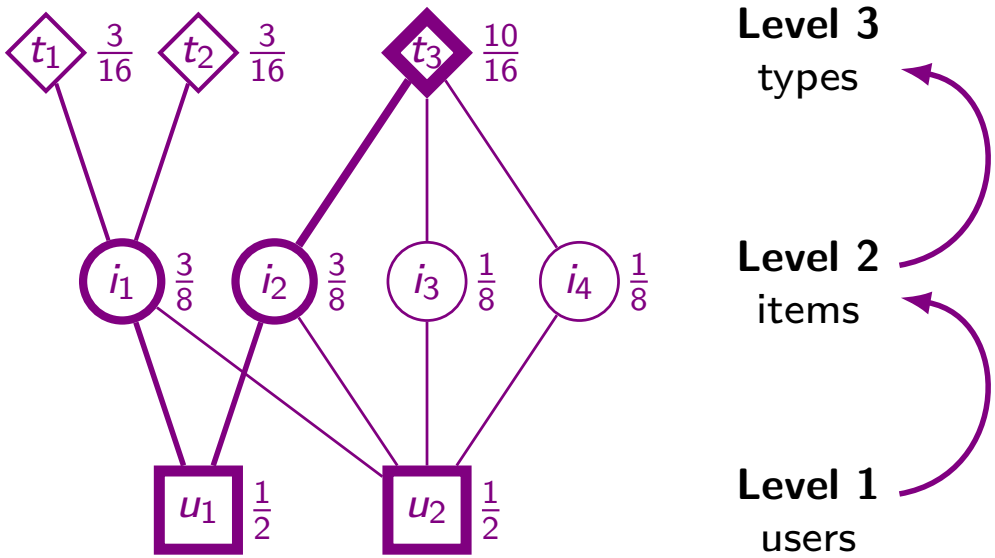
	Mean of individual ...	Collective ...
... item-richness of users	17.5402002	5327
... type-richness of users	9.0708963	45

Collective type-diversity of items (uniform initial distribution)



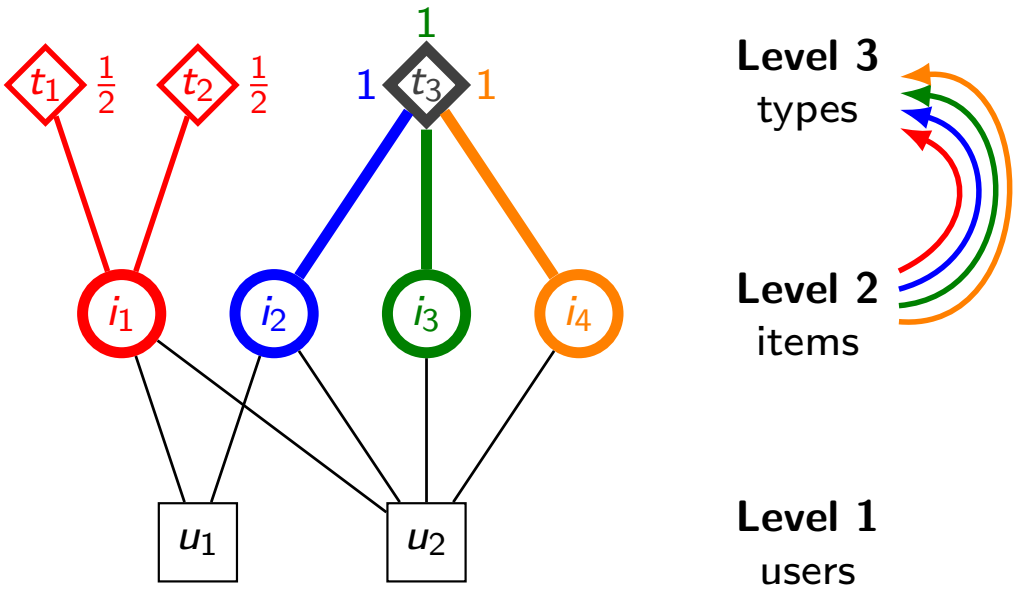

```
get_diversity_from_path (
    graph = example,
    path = c(2,3),
    type = 'collective',
    measure = c('richness', 'entropy', 'herfindahl')
)

##   richness    entropy herfindahl
##   3.000000   1.061278   0.593750
```

Collective type-diversity of items (biased initial distribution)



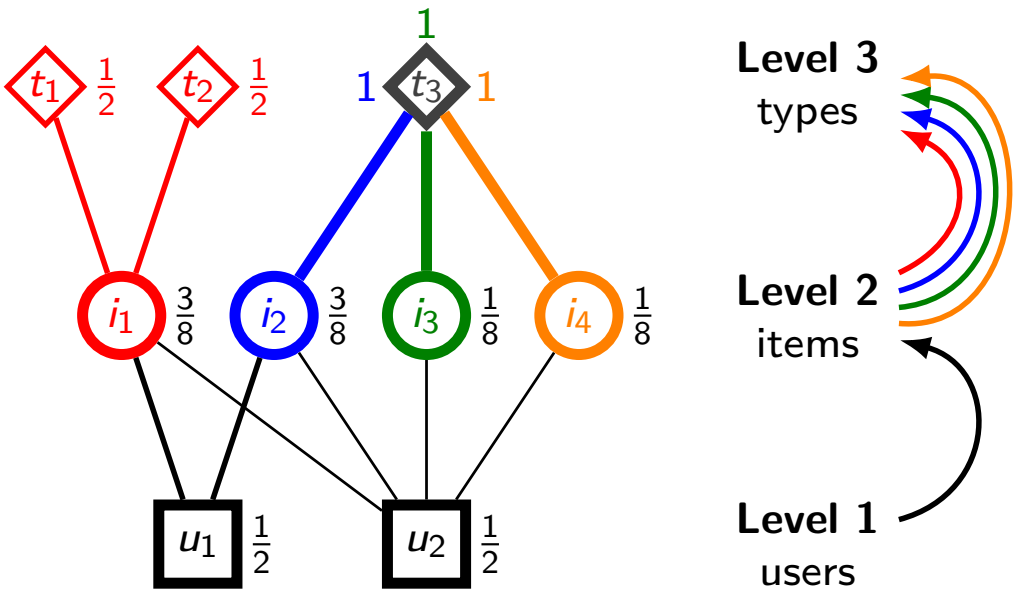

```
get_diversity_from_path (
    graph = example, path = c(1,2,3),
    type = 'collective',
    measure = c('richness', 'entropy', 'herfindahl')
)

##   richness    entropy herfindahl
##  3.0000000  1.3294340  0.4609375
```

Mean of individual type-diversity of items (uniform mean)



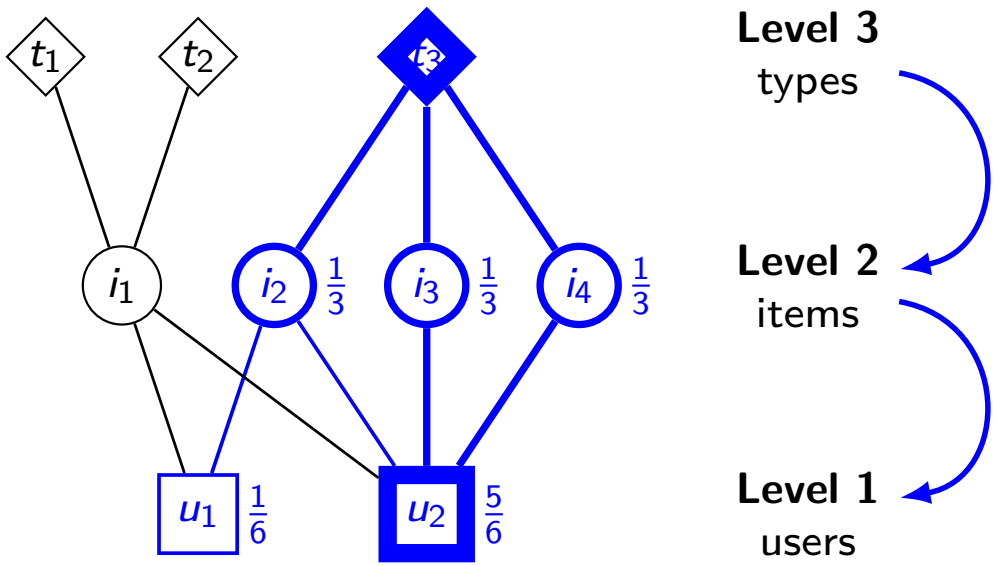

```
get_diversity_from_path (
    graph = example, path = c(2,3),
    type = 'mean',
    measure = c('richness', 'entropy', 'herfindahl')
)


##   richness    entropy herfindahl
##  1.1892071  0.2500000  0.8408964
```

Mean of individual type-diversity of items (biased mean)




```
get_diversity_from_path (
    graph = example, path = c(2,3),
    type = 'mean',
    mean_distribution = get_distribution_from_path (example, c(1,2)),
    measure = c('richness', 'entropy', 'herfindahl')
)

##   richness    entropy herfindahl
##  1.2968396  0.3750000  0.7711054
```

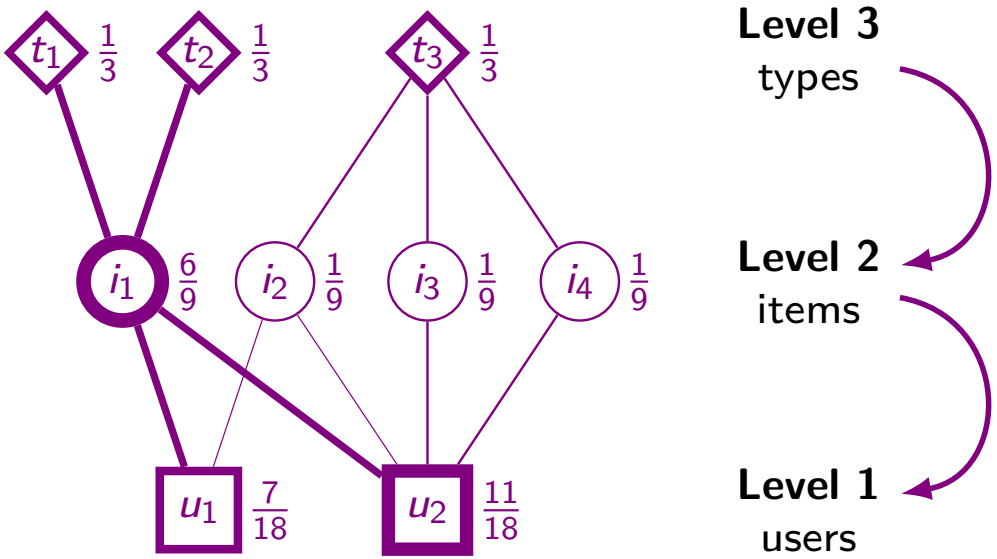
Individual user-diversity of a given item




```
get_diversity_from_path (
    graph = example, path = c(3,2,1),
    initial_node = 't3',
    measure = c('richness', 'entropy', 'herfindahl')
)
```

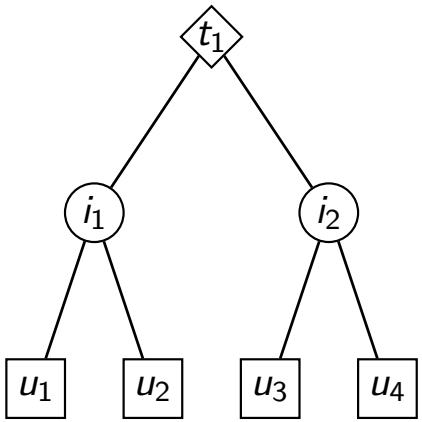
```
##   richness    entropy herfindahl
##  2.0000000  0.6500224  0.7222222
```

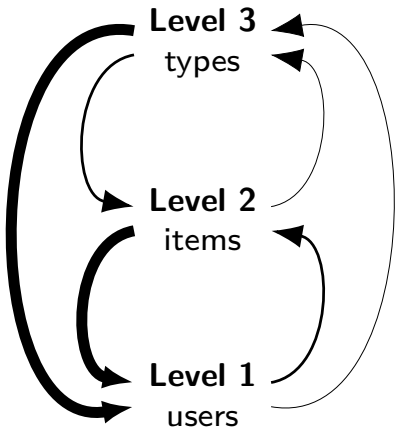
Collective user-diversity of items

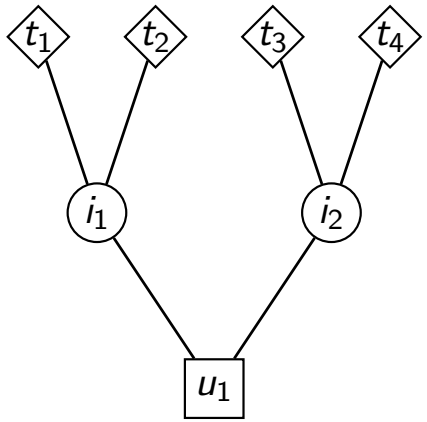


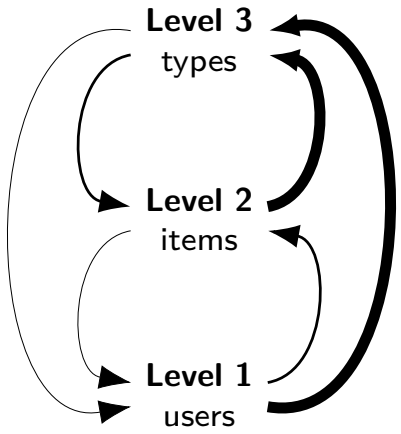

```
get_diversity_from_path (
    graph = example, path = c(3,2,1),
    type = 'collective',
    measure = c('richness', 'entropy', 'herfindahl')
)

##   richness    entropy herfindahl
##  2.0000000  0.9640788  0.5246914
```

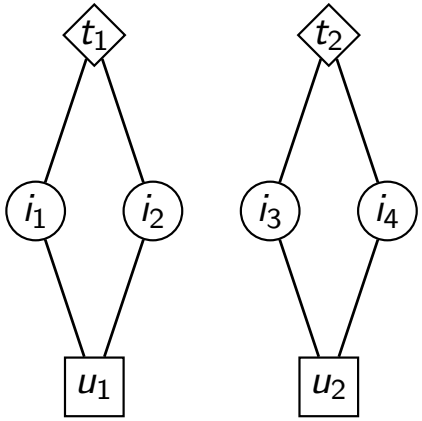
Different paths for different diversity patterns

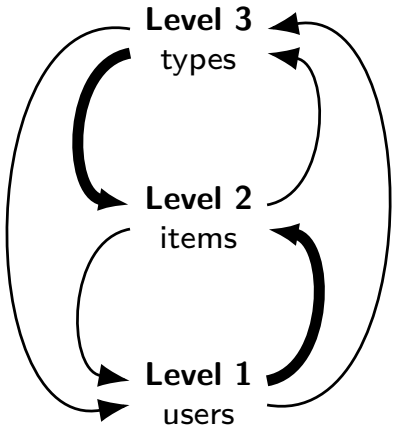


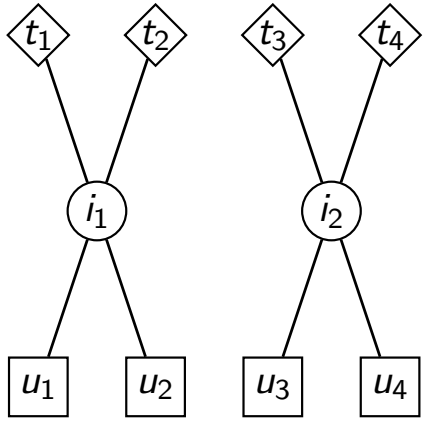


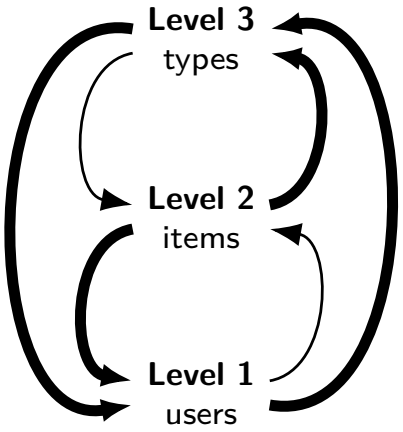






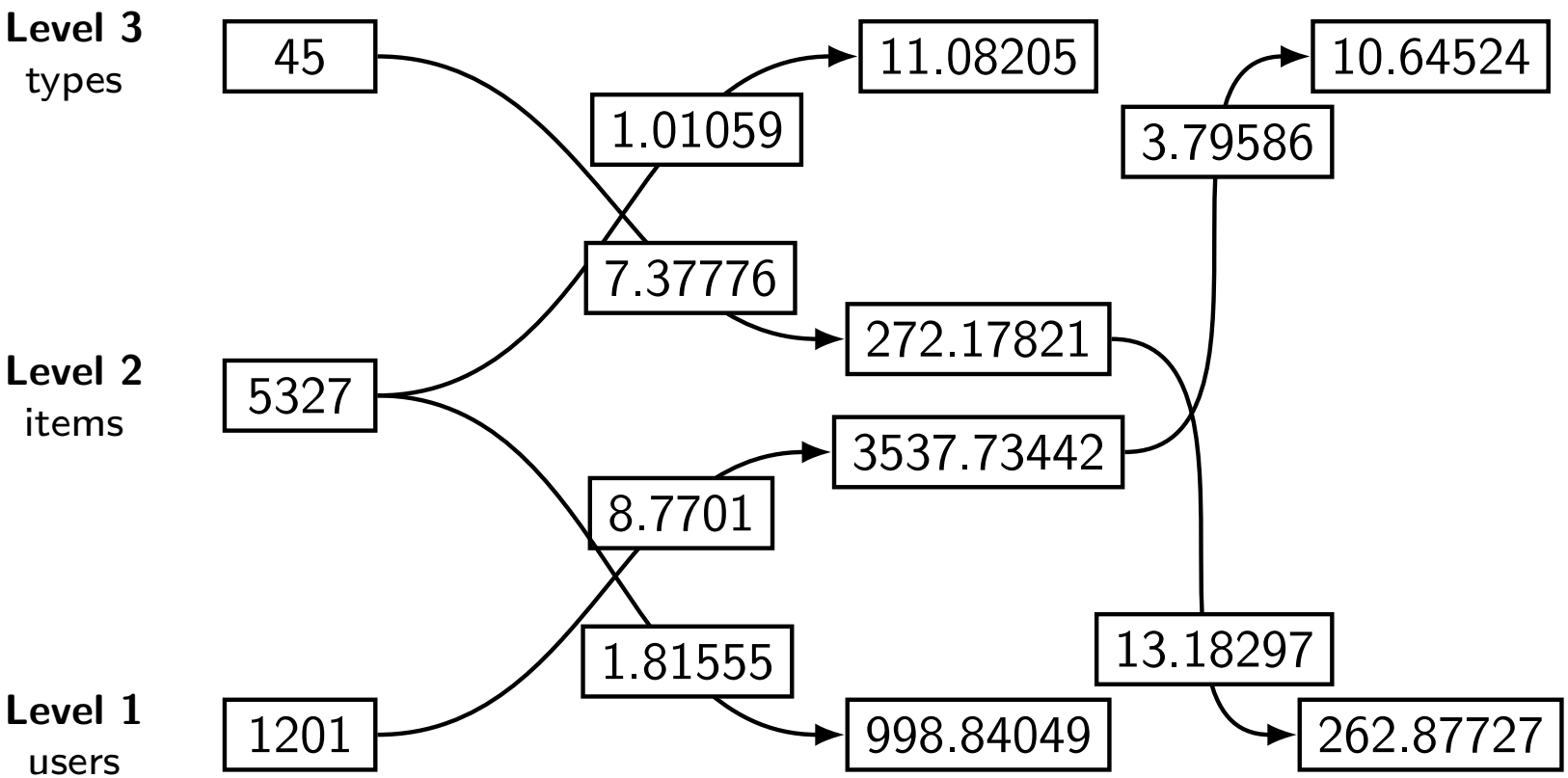




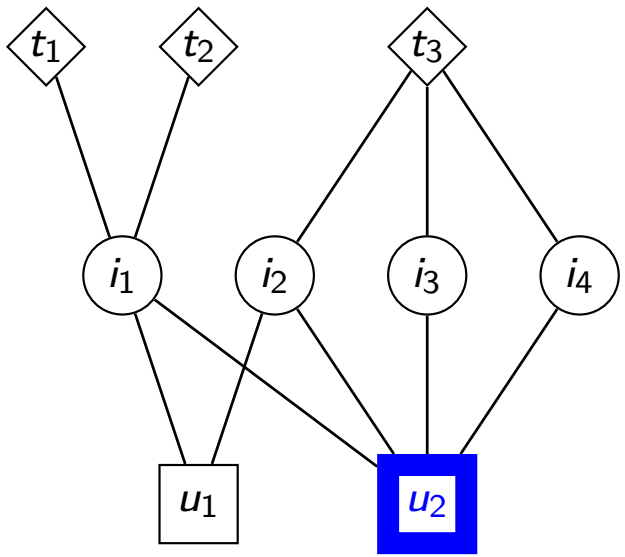


Diversity diagram

```
get_all_diversities (automotive, length=2, cycles=FALSE, order=1)
```



Cycling user-diversity of a given user



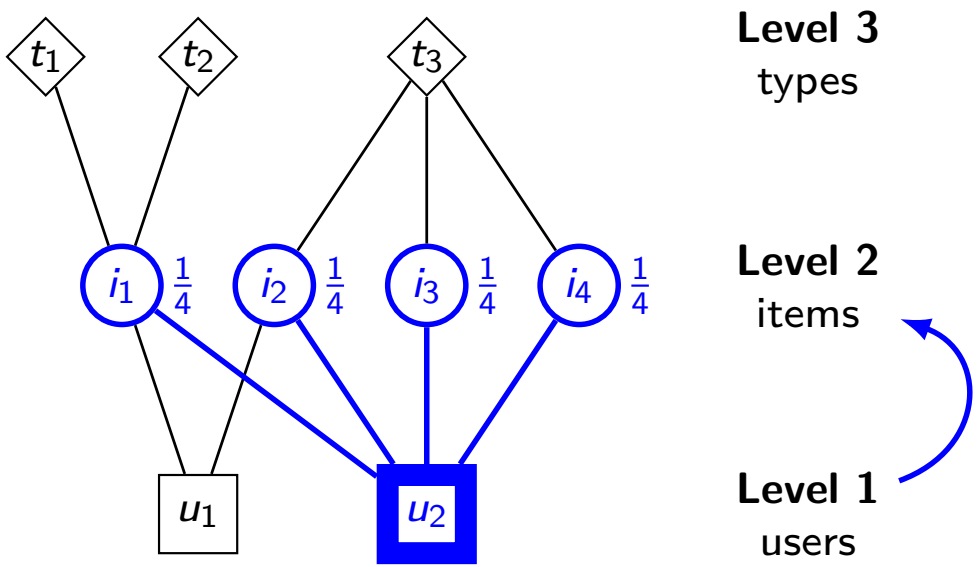
Level 3
types

Level 2
items

Level 1
users

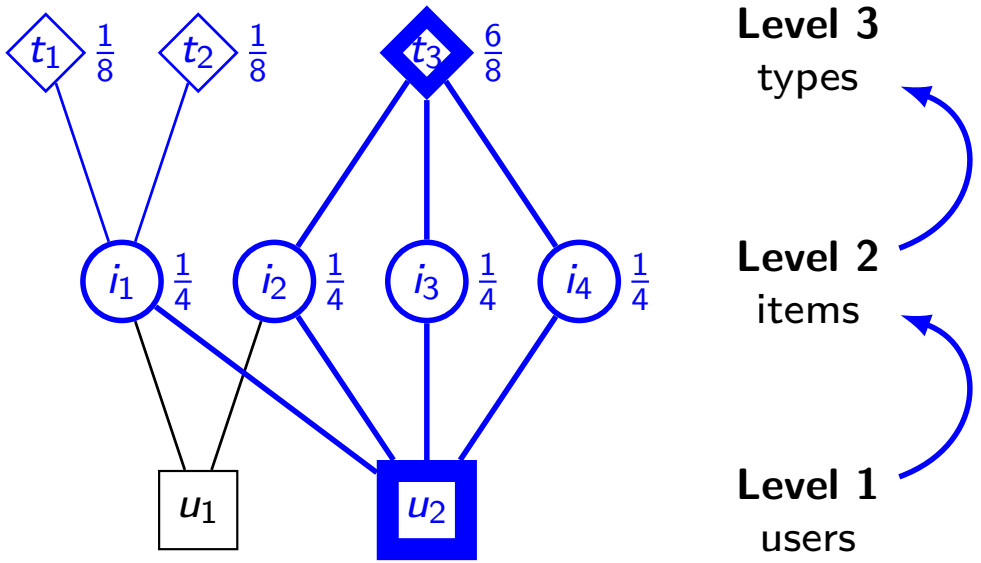
```
get_diversity_from_path (
    graph = example, path = c(1,2,3,2,1),
    initial_node = 'u2',
    measure = c('richness', 'entropy', 'herfindahl')
)

##   richness    entropy herfindahl
##  2.0000000  0.8112781  0.6250000
```

Cycling user-diversity of a given user



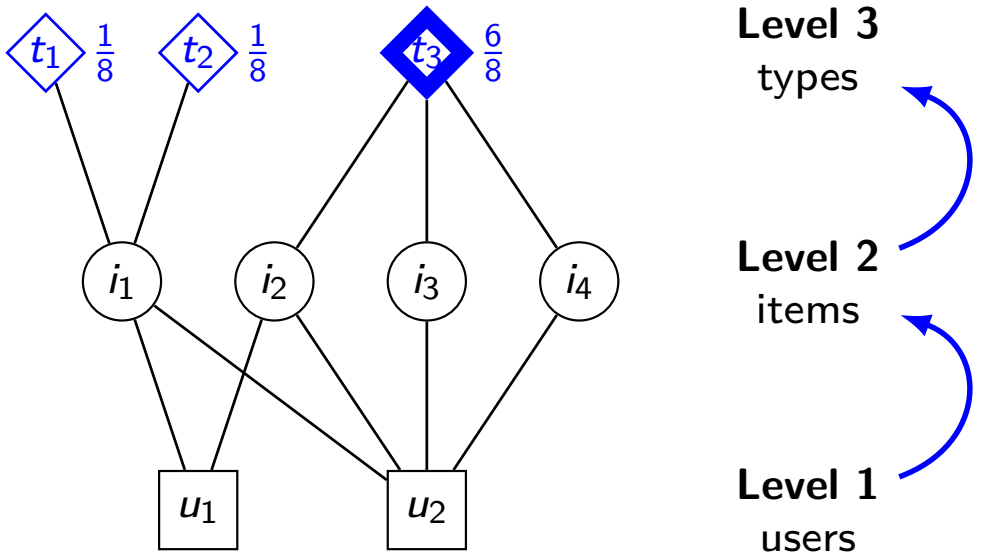

```
get_diversity_from_path (
    graph = example, path = c(1,2,3,2,1),
    initial_node = 'u2',
    measure = c('richness', 'entropy', 'herfindahl')
)

##   richness    entropy herfindahl
##  2.0000000  0.8112781  0.6250000
```

Cycling user-diversity of a given user



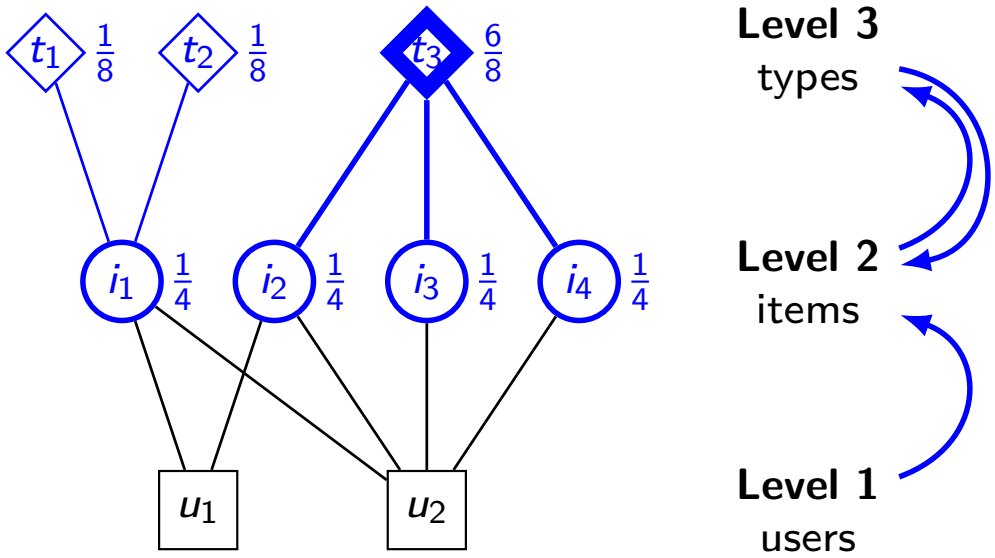

```
get_diversity_from_path (
    graph = example, path = c(1,2,3,2,1),
    initial_node = 'u2',
    measure = c('richness', 'entropy', 'herfindahl')
)

##   richness    entropy herfindahl
##  2.0000000  0.8112781  0.6250000
```

Cycling user-diversity of a given user



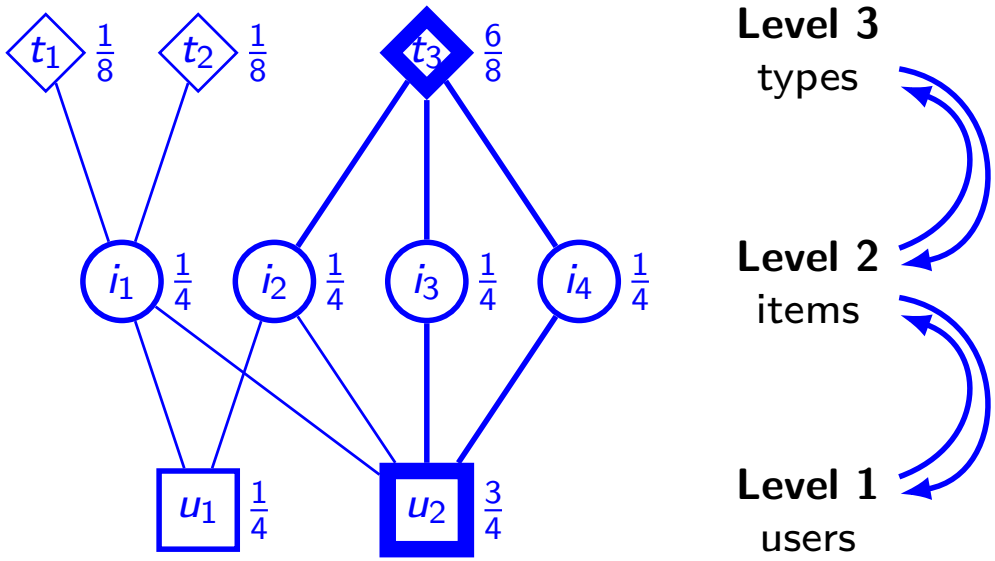

```
get_diversity_from_path (
    graph = example, path = c(1,2,3,2,1),
    initial_node = 'u2',
    measure = c('richness', 'entropy', 'herfindahl')
)

##   richness    entropy herfindahl
##  2.0000000  0.8112781  0.6250000
```

Cycling user-diversity of a given user



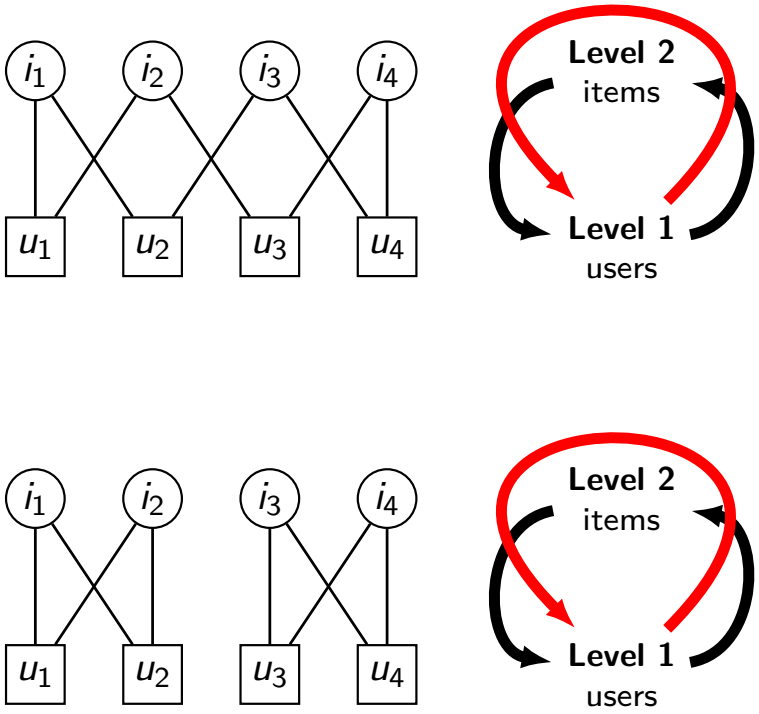

```
get_diversity_from_path (
    graph = example, path = c(1,2,3,2,1),
    initial_node = 'u2',
    measure = c('richness', 'entropy', 'herfindahl')
)

##   richness    entropy herfindahl
##  2.0000000  0.8112781  0.6250000
```

Cycling user-diversity of a given user




```
get_diversity_from_path (
    graph = example, path = c(1,2,3,2,1),
    initial_node = 'u2',
    measure = c('richness', 'entropy', 'herfindahl')
)

##   richness    entropy herfindahl
##  2.0000000  0.8112781  0.6250000
```

Cycling diversity vs. Collective and Individual diversities

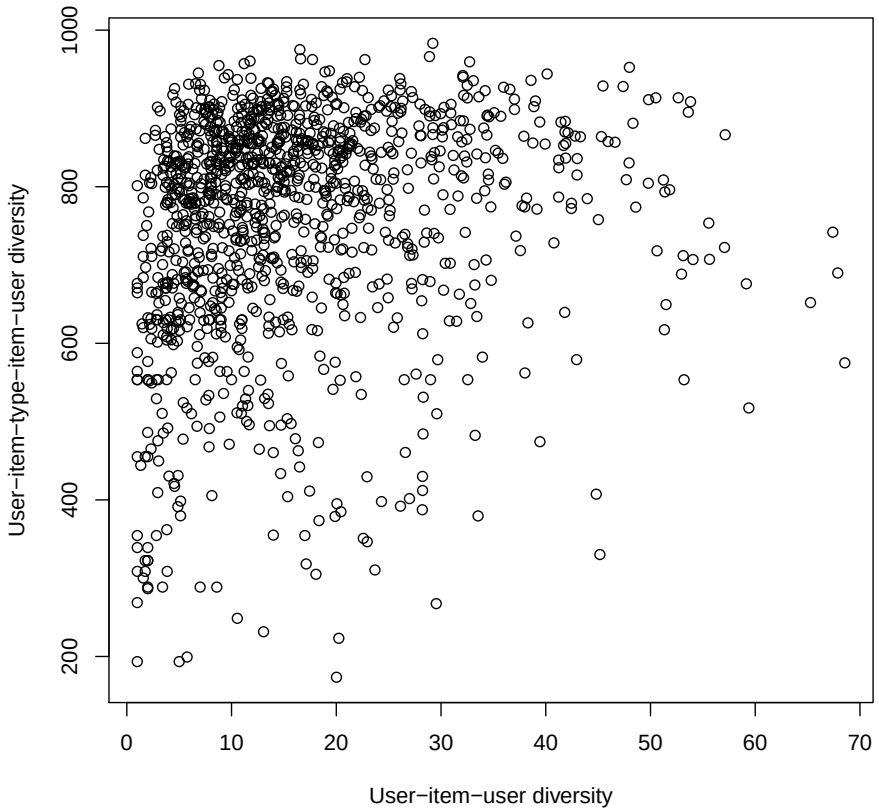


Same collective diversity
Same individual diversity

Different cycling diversity!

User-item-user vs. User-item-type-item-user diversity

```
user_item_user_div <- get_diversity_from_path (
    graph=automotive, path=c(1,2,1), order=1)
user_type_user_div <- get_diversity_from_path (
    graph=automotive, path=c(1,2,3,2,1), order=1)
plot (user_item_user_div, user_type_user_div,
      xlab='User-item-user diversity', ylab='User-item-type-item-user diversity')
```



The End
Thanks for your attention