

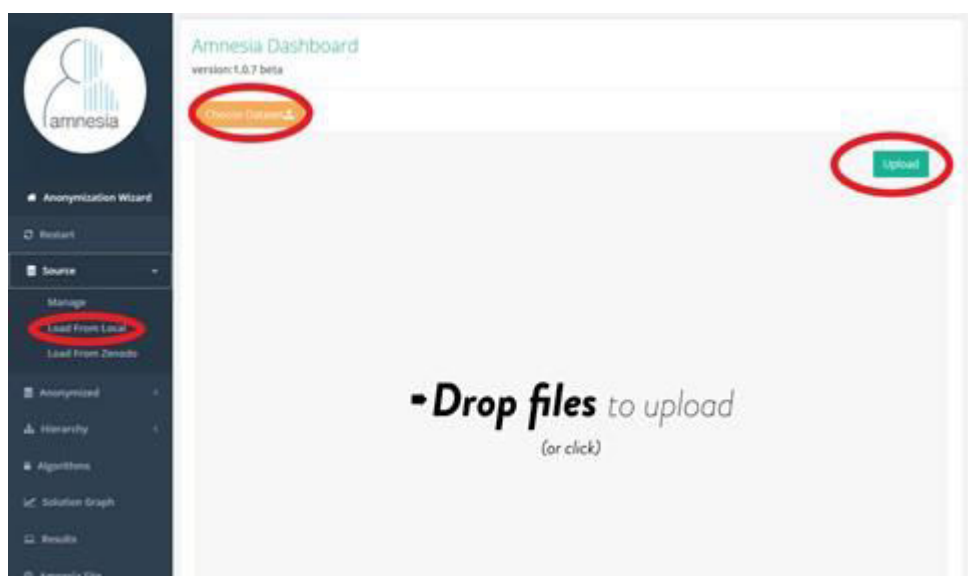
K^m – anonimnost

K^m – anonimnost je slabiji oblik k – anonimnosti koji je pogodniji za višedimenzionalne lične podatke. Kao i kod k – anonimnosti, algoritam uzima u obzir broj n kvazi identifikatora, ali sada ograničava garanciju privatnosti samo na m (bilo koji m) od n kvazi identifikatora. Drugim rečima, algoritam anonimizacije garantuje da će se svaka kombinacija od m kvazi identifikatora pojaviti k puta u anonimiziranim skupovima podataka, nezavisno od ukupnog broja n kvazi identifikatora. Tipično, $m \ll n$. K^m – anonimnost se koristi za skupove podataka sa skupom postavljenih vrednosti atributa u Amnesia aplikaciji (<https://amnesia.openaire.eu/>).

Skup postavljenih vrednosti (eng. *set-values*)

Objektno–relacioni skup podataka (eng. *dataset*) je kombinacija relacionog skupa podataka i skupa postavljenih vrednosti atributa koji se odnose na skup podataka. **Objektno–relacioni** skup podataka sadrži proizvoljan broj kolona sa jednostavnim tipovima vrednosti i samo jedan atribut iz skupa vrednosti. Pojedine vrednosti u skupu postavljenih vrednosti atributa su odvojene graničnikom, različitim od graničnika koji označava kolone (attribute). Skupovi podataka kojima upravlja Amnesia sadrže najviše jedan atribut sa skupom postavljenih vrednosti. Jednostavne vrednosti mogu biti samo stringovi. Amnesia primenjuje samo K^m – anonimnost na objektno-relacionim skupovima podataka koristeći **apriori** algoritam koji daje jedno rešenje sa lokalno optimalnim „rezom“ u hijerarhiji generalizacije. Sledeći primer pokazuje kako primeniti K^m – anonimnost u objektno-relacionom skupu podataka.

Prvo, korisnik bira odgovarajuću datoteku iz svog računara.



Korisnik mora odabrati tip skupa podataka tj. “Relational and Set” za ovaj primer. Konačno, korisnik mora definisati dva graničnika: a) jedan za različite kolone (attribute) i b) poseban za rasčlanjavanje pojedinačnih vrednosti unutar atributa sa skupom postavljenih vrednosti.

This is how the dataset looks like :

Procedure Codes
926 716 671 829
30 488 860 976
328

...

Delimiter * delimiter (in the example is " | ")

Data Set Type : dataset type

Amnesia će predložiti vrstu podataka, ali korisnik ima priliku da promeni procene ove aplikacije. Atribut sa skupom postavljenih vrednosti biće posebnog tipa “set”, što podrazumeva skup stringova.

What type is your data?

Choose the columns and their types.

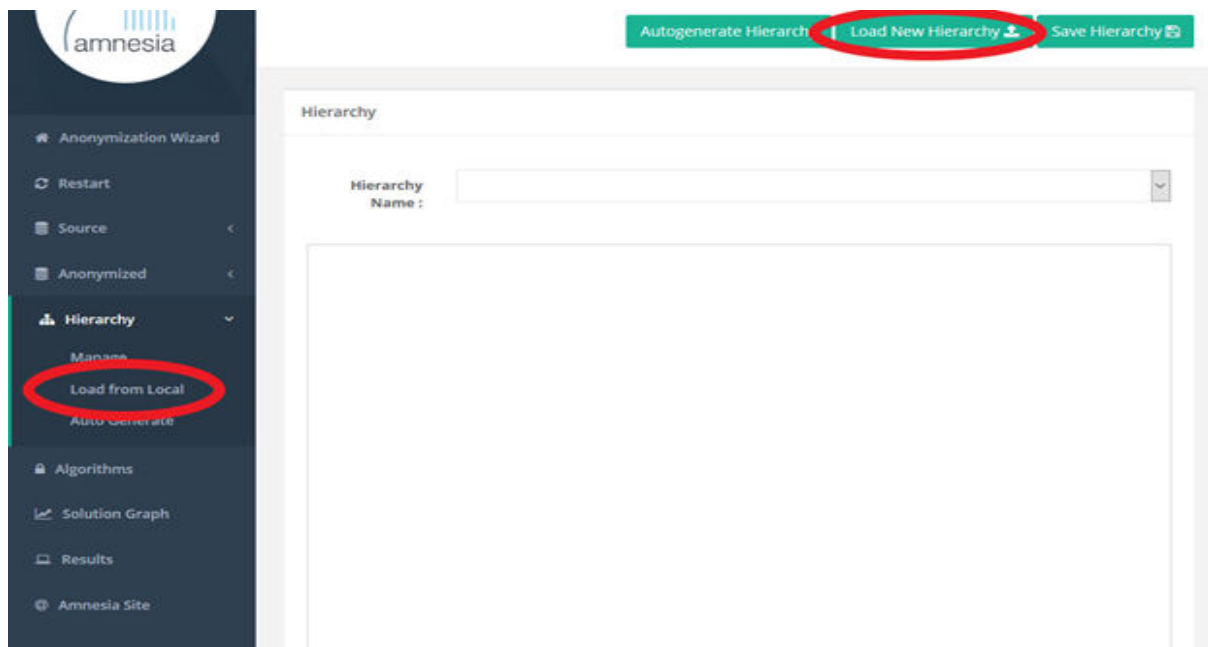
Procedure Codes
926 716 671 829
30 488 860 976
328
786 496
273 365 202 141 522

Predstavljanje hijerarhije generalizacije

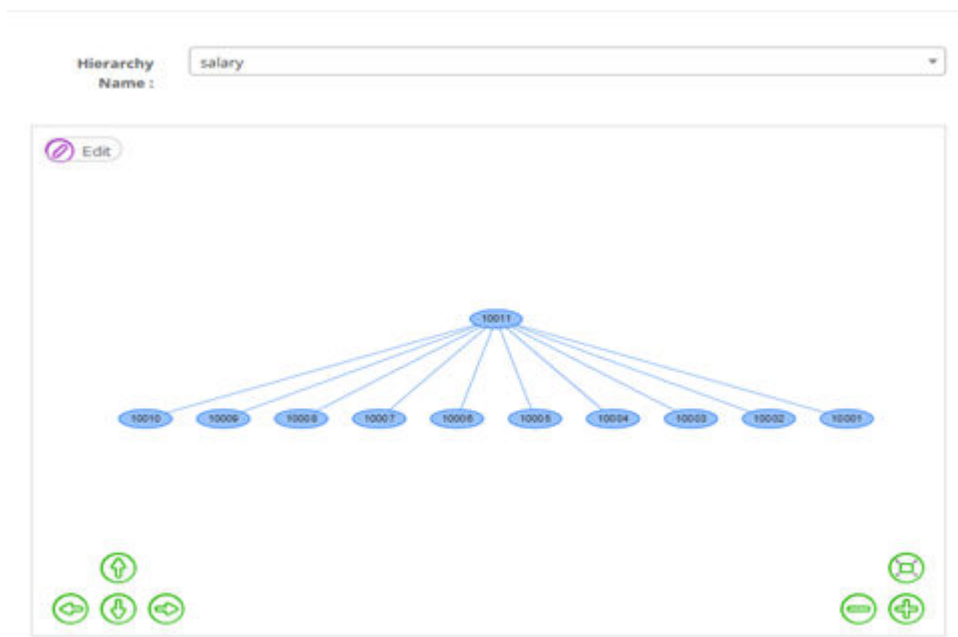
Algoritam ove aplikacije k – anonimnosti koristi generalizaciju za stvaranje grupa identičnih podataka. Jedinstvene ili retke vrednosti (pojavljuju se manje od k puta u skupu podataka) ili kombinacije vrednosti zamenjuju se opštijim vrednostima. Na primer, ako u skupu podataka postoji samo jedna osoba koja ima prebivalište u Grčkoj, tada će Grčka (i ostale zemlje EU) biti zamenjene sa EU kako bi se stvorilo više od k identičnih vrednosti u atributu Države prebivališta (“Country of Residence”). Korisnik zadaje hijerarhiju generalizacije koja definiše kako se odvijaju ove zamene, kao ulaz u algoritam.

Amnesia može da pomogne korisniku da kreira hijerarhije koristeći postojeće skupove podataka, više o tome kako kreirati novu i urediti postojeću hijerarhiju, može se naći u **Tutorial**.

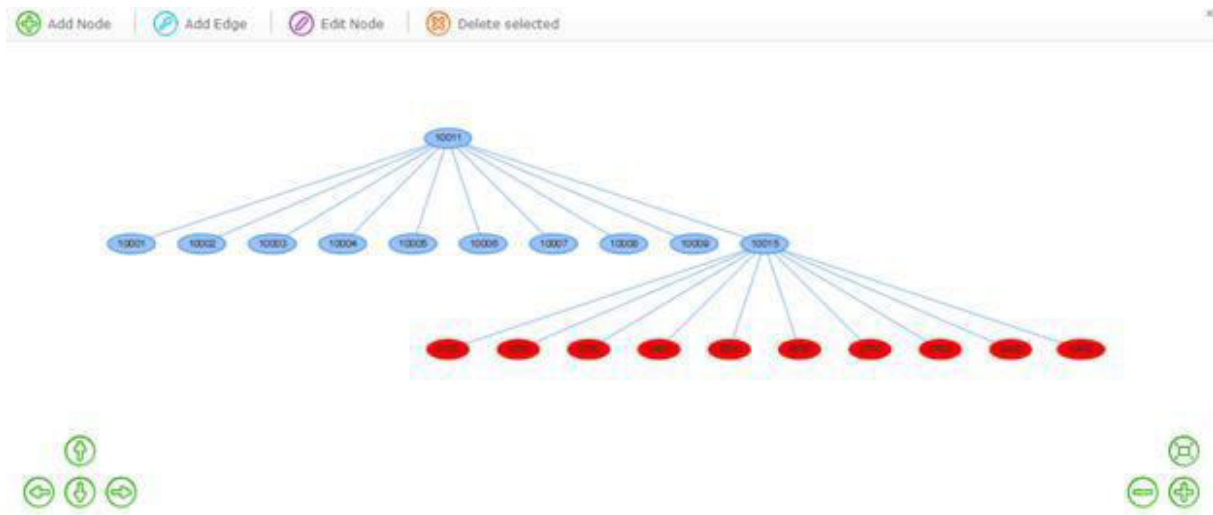
Korisnik može učitati hijerarhiju na dva različita načina: a) klikom na dugme “Load New Hierarchy” ili b) klikom na dugme “Load from Local” u levom bočnom meniju.



Učitane hijerarhije se na ekranu pojavljuju u vidu grafikona, gde određeni čvorovi ukazuju na to kako se generišu njegovi podređeni čvorovi.



Dvostrukim klikom na čvor mogu se videti podređeni čvorovi.



Izvršenje algoritama i rezultati

Nakon što se skup podataka i hijerarhije učitaju, korisnik poziva algoritam za anonimizaciju. Potrebno je da korisnik: a) poveže svaki atribut sa hijerarhijom koja se na njemu koristi i b) definiše vrednosti k i m . Dugme za izvršenje “Execute” pokreće algoritam za anonimizaciju.

Showing 1 to 10 of 999 entries

Procedure Codes
926 716 671 829
30 488 860 976
328
786 496
273 365 202 141 522
318 476 783 917
383 392 399
796 740 730 892 398
875 255 433 479
17 533 786 433 91

Showing 1 to 10 of 999 entries

Previous 1 2 3 4 5 ... 100 Next

Bind Hierarchies with Attributes

Indicate with generalization hierarchy will be used for each dataset attribute. The same hierarchy can be used in multiple attributes. A hierarchy must be defined for each quasi identifier.

Procedure Codes myNumHier

Algorithms

Type: Apriori K: 3 M: 2 Execute

Algoritam je potom spreman za izvršavanje. Nakon izvršenja anonimizirani i originalni skup podataka uporedo se prikazuju na ekranu. Korisnik može preuzeti anonimizirani skup podataka ili pravila anonimizacije.

DataSet

Show10entries

original dataset

Procedure Codes

926|716|671|829

30|488|860|976

328

786|496

273|365|202|141|522

318|476|783|917

383|392|399

796|740|730|892|398

875|255|433|479

17|533|786|433|91

Showing 1 to 10 of 999 entries

Previous

1

2

3

4

5

...

100

Next

Anonymized DataSet

anonymized dataset

Procedure Codes

Random104,Random102,Random103

Random100,Random98,Random104,Random105

Random98

Random100,Random103

Random100,Random98,Random97,Random96

Random100,Random98,Random104,Random103

Random99

Random99,Random104,Random102,Random103

Random100,Random99,Random97,Random104

Random100,Random99,Random96,Random104,Random103

Statistics

NAPOMENA: Svi termini koji su korišćeni u ovom dokumentu i koji su izraženi u gramatičkom muškom rodu, porazumevaju prirodni ženski i muški rod.

Prevela dokument „Amnesia Anonymity Tutorial“, pod CC licencom, (<https://amnesia.openaire.eu/Scenarios/AmnesiaKMANonymityTutorial.pdf>) sa engleskog na srpski Jelena Nikolić u sklopu volontiranja za Serbia.RDA grupu u Beogradu aprila 2021.