Detekcija zajednica na socijalnim mrezama korišćenjem genetskog algoritma

- Projekat u okviru predmeta Racunarska inteligencija -

Anđela Đurović

26/2019

Sadržaj

[**1.** **Opis problema** 3](#_Toc146711407)

[**2.** **Implemetacija** 3](#_Toc146711408)

[**2.1. Ulazni podaci** 3](#_Toc146711409)

[**2.2. Implementacija jedinke** 3](#_Toc146711410)

[**2.3. Implementacija selekcije, ukrstanja i mutacije** 4](#_Toc146711411)

[**2.4. Parametri genetskog algoritma** 4](#_Toc146711412)

[**3. Rezultati** 4](#_Toc146711413)

[**4. Reference** 5](#_Toc146711414)

# **Opis problema**

Za dati skup tačaka nodes, i veze između tačaka edges, potrebno je da identifikujemo gusto povezane grupe čvorova sa retkim vezama između grupa.

# **2.** **Implemetacija**

## **2.1. Ulazni podaci**

Ulazne podatke učitavamo tako sto imamo zadate čvorove i grane izmedju čvorova (neusmeren graf)

## **2.2. Implementacija jedinke**

Populacija od sastoji od jediniki, a jedinku pravimo tako što uzimamo na slučajan način za svaki čvor neki njegov sused.

*Fitness* funkcija je zadata kao:

Gde je , dok je , pri čemu je definisano kao

## **2.3. Implementacija selekcije, ukrstanja i mutacije**

Selekcija je implementirana koriscenjem rulet selekcije, gde svakom roditelju dodeljujemo težine kojima ih biramo.

Korišćeno je višepoziciono ukrštanje. Tako što na slučajan način uzimamo jednu vrednost iz jednog ili drugog roditelja. Na ovaj način smo postigli da većih promena jedinki.

Mutacija se odnosi na promenu vrednosti suseda jednog čvora, tako što ćemo zameniti sa nekim drugim susedom tog čvora.

**2.4. Parametri genetskog algoritma**

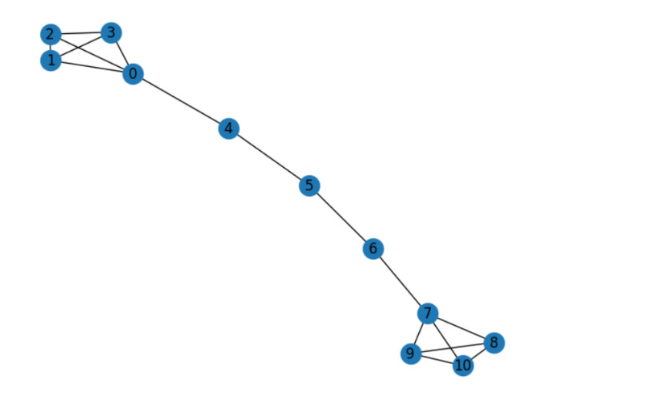
Vrednosti za parametre genetskog algoritma su uglavnom empirijski određene. Za veličinu populacije odabrano je 200 jedinki, za broj generacije je uzeto 30, dok elitizmom čuvamo najboljih 10% jedinki.

# **3. Rezultati**

U narednom delu prikazani su rezultati koje smo dobili na osnovu sledećeg *inputa*:

nodes = [0,1,2,3,4,5,6,7,8,9,10]

edges= [(0,1),(0,4),(1,2),(2,3),(1,3),(3,0),(0,2),(4,5),(5,6),(6,7),(10,8),(10,9),(8,9),(8,7),(9,7),(7,10)]



Genetskim algoritmom detektovani su sledeci klasteri:

[[0, 1, 2, 3], [4, 5, 6], [8, 9, 10, 7]],

dok je brute-force algoritam detektovao:

[ [0, 1, 2, 3] ] , [ [4, 5, 6, 7, 8, 9, 10] ]

Da bi brute force uspeo da detektuje vise klasa potrebno je dosta izmena u kodu I primene razlicitih algoritama, od kojih je jedan od jednostavnijih I veoma efikasnih Genetski algoritam.

# **4. Reference**

* [A Genetic Algorithm for Community Detection in Social Networks](https://staff.icar.cnr.it/pizzuti/pubblicazioni/PPSN08.pdf)
* Angiulli, F., Cesario, E., Pizzuti, C.: A greedy search approach to co-clustering sparse binary matrices. In: Proceedings of the 18th International Conference on Tools with Artificial Intelligence (ICTAI 2006), pp. 363–370 (2006)