**Übungsblatt 10**  
Systembiologie  
Deadline: 13.7.2022 um 20:00 MEZ  
Bioinformatik für Biochemiestudierende  
Dr. Florian Klimm  
Sommersemester 2022

**Aufgabe 10.1 Netzwerkanalyse (40 %)**  
Zeichnen Sie einen Graphen mit der folgenden Edgelist: (1 , 2),(2, 4),(2 , 3),(2, 5),(5 , 6),(3, 7),(6 , 7)

1

2

4

3

6

5

7

Berechnen sie zusätzlich den Degree, die Betweenness, die Closeness, und den  
Clusteringkoeffizienten aller Knoten.

**Adjazanzmatrix:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1 | **0** | **1** | **0** | **0** | **0** | **0** | **0** |
| 2 | **1** | **0** | **1** | **1** | **1** | **0** | **0** |
| 3 | **0** | **1** | **0** | **0** | **0** | **0** | **1** |
| 4 | **0** | **1** | **0** | **0** | **0** | **0** | **0** |
| 5 | **0** | **1** | **0** | **0** | **0** | **1** | **0** |
| 6 | **0** | **0** | **0** | **0** | **1** | **0** | **1** |
| 7 | **0** | **0** | **1** | **0** | **0** | **1** | **0** |

**Mean degree:**

With and

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Node i | 1 | 2 | 3 | 4 | 5 | 6 | 7 |  |  |
| Degree ki | 1 | 4 | 2 | 1 | 2 | 2 | 2 | 14 | 14/7=2 |

**Clustering coefficient:**

Hier ist M die Anzahl an edges zwischen den Knoten, mit denen der Knoten verknüpft ist, den man sich anschauen möchte.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Node i | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Degree ki | 1 | 4 | 2 | 1 | 2 | 2 | 2 |
| Coefficient Ci |  |  |  |  |  |  |  |

**Closeness:**

Distance Matrix

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1 | **0** | **1** | **2** | **2** | **2** | **3** | **3** |
| 2 | **1** | **0** | **1** | **1** | **1** | **2** | **2** |
| 3 | **2** | **1** | **0** | **2** | **2** | **2** | **1** |
| 4 | **2** | **1** | **2** | **0** | **2** | **3** | **3** |
| 5 | **2** | **1** | **2** | **2** | **0** | **1** | **2** |
| 6 | **3** | **2** | **2** | **3** | **1** | **0** | **1** |
| 7 | **3** | **2** | **1** | **3** | **2** | **1** | **0** |

1. D=13
2. D=8
3. D=10
4. D=13
5. D=10
6. D=12
7. D=12

**Betweenness:**

Wieviele Komponenten (connected components) hat der Graph? Berechnen sie  
dessen Dichte (density).

Der Graph hat eine Komponente.

**Aufgabe 10.2 Netzwerkanalyse in R (40 %)**  
Nutzen sie die R Library iGraph um ein Proteininteraktionsnetzwerk zu untersu-  
chen. Sie finden Informationen über das Paket unter https://igraph.org/r/  
und https://kateto.net/networks-r-igraph.  
Führen Sie dazu die folgenden Schritte durch  
1. Laden der Edgelist aus proteinInteractionNetworkHomoSapiensSmall.csv  
2. Einen Graphen daraus erstellen  
3. Den Graphen Illustrieren  
4. Berechnen des Degrees, der Closeness, und der Betweenness für alle Knoten  
5. Bestimmen der wichtigsten Knoten im Graphen  
Diskutieren Sie Ihre Ergebnisse.

**Aufgabe 10.3 Freundschaftsparadox (20 %)**  
Lesen Sie den Artikel Friends You Can Count On von Steven Strogatz (verfügbar un-  
ter https://opinionator.blogs.nytimes.com/2012/09/17/friends-you-can-count-on/).  
Beschreiben Sie kurz, in eigenen Worten, was das Freundschaftsparadox bedeutet.  
Inwieweit kann das Freundschaftsparadox in der Biologie genutzt werden?

Das Freundschaftsparadox bedeutet folgendes: Schaut man sich an, wie viele Freunde ein Individuum hat und vergleicht dies mit der Anzahl von Freuden, die eben diese Freunde besitzen, so ist die Anzahl der Freunde von den Freunden erstaunlicherweise immer größer als die Anzahl der Freunde des Individuums. Das Freundschaftsparadox lässt sich mathematisch erklären. Grob gesagt steckt hinter dem Paradox die Tatsache, dass die Anzahl von Freunden von Freunden durch den *wheigted average* berechnet wird. Hier wird dann praktisch aus der Sicht jedes Freundes geschaut, wie viele Freunde die anderen Personen jeweils haben, sodass der Score jedes Individuums mit sich selbst multipliziert wird. Das Quadrieren gewichtet somit Freunde mit vielen Freunden besonders stark und erhöht des *weighted average.*

Eine Anwendung ist die frühe Detektion von Ausbrüchen von Infektionskrankheiten. Da es sehr schwierig ist, ein ganzes soziales Netzwerk tatsächlich zu „mappen“, um herauszufinden, welche Individuen am stärksten vernetzt sind, können die Freunde zufällig selektierter Individuun überwacht werden, da diese dem Paradox zufolge vermutlich besser vernetzt sind. Durch diese Monitoring-Methode könnte erheblich an Zeit gewonnen werden bei der Reaktion auf Ausbrüche von infektiösen Krankheiten.