

BLM2512 Proje Ödevi: Sosyal Ağ Analizi

Sosyal Ağ Analizi, sosyolojide graflar gibi analitik araçları kullanarak sosyal oluşumları araştıran bir çalışma alanıdır. Bu proje ödevinde, basit bir sosyal ağ analizi algoritmasını C dilinde gerçekleştirmeniz beklenmektedir.

(GitHub'dan alınan) Veri CSV formatında sunulmuştur ve ilk satırları şu şekildedir:

follower,followed

1570,9236

9236,1570

13256,9236

...

İlk satırda 1570 ID'li kullanıcı, 9236 ID'li kullanıcıyı takip etmektedir.

İkinci satırda, 9236 ID'li kullanıcı, 1570 ID'li kullanıcıyı takip etmektedir.

Üçüncü satırda, 13256 ID'li kullanıcı, 9236 ID'li kullanıcıyı takip etmektedir.

Yapılması Beklenenler:

1. Grafi okumak ve yüklemek:

Dosya adını girdi olarak alan ve CSV formatında okuduğu verileri bir komşuluk listesi (adjacency list) veri yapısına yükleyen bir fonksiyon yazınız. **(3 puan)**

2. Temel graf metrikleri:

En çok bağlantısı olan 10 kullanıcı hangileridir? **(4 puan)**

Tüm bağlantı sayılarının ortalama değeri ve medyan değeri nedir? **(3 puan)**

3. En kısa yollar:

En kısa yollar, pek çok ağ ölçütü için temel oluşturur. Dijkstra'nın algoritmasını gerçekleyebilirsiniz.

Tüm düğüm çiftleri arasındaki en kısa yolları hesaplayan bir fonksiyon yazınız. **(7 puan)**

4. Önemli kullanıcıları derecelendirme:

Sosyal ağ analizinin temel kullanımlarından biri, önemli / etkili düğümleri belirlemektir. Bir düğümün önemini ölçmek için kullandığımız tipik bir ölçü merkeziliyet (centrality). Merkeziliyeti hesaplamak için literatürde pek çok yöntem mevcuttur ancak bizim için arada bulunma ölçütünü (betweenness centrality measure) hesaplamak yeterlidir.

İnternette daha detaylı bilgi bulabilirsiniz ancak basit bir sözde-kod (pseudo-code) şu şekildedir:

Bir D düğümünün arada bulunma ölçütünü hesaplamak için:

1. Tüm (d1, d2) düğüm çiftleri için, aralarındaki en kısa yolları hesapla.
2. Tüm (d1, d2) düğüm çiftleri için, D düğümünü içeren en kısa yolların oranını hesapla.
3. (d1, d2) düğüm çiftlerinin tümünde bu oranı topla.

Verilen sosyal ağ üzerinde tüm düğümlerin arada bulunma merkeziliyet ölçütünü hesaplayan bir fonksiyon yazınız. Bu fonksiyon ile kullanıcıları derecelendiriniz ve en önemli 10 kullanıcıyı, önem değerleri ile birlikte listeleyiniz. **(13 puan)**

Teslim edilecekler: TÜMÜNÜ TEK BİR SIKIŞTIRILMIŞ DOSYA HALİNDE TESLİM EDİNİZ !!!

1. C kodunuz.

* Kod içinde fonksiyonlarınızın başında, algoritmik çözüm yaklaşımınızı yorum satırları içinde anlatınız.

2. Yapılması beklenen 4 adımın sonuçlarını ve çıktılarını içeren bir rapor.

* Word/PDF vb şekilde hazırlayabilirsiniz.

Son Teslim Tarihi: 14.01.2021, 11.59

Toplam Not Değeri: 30 puan