**ENF-203**

**FİNAL CEVAPLARI**

1. Düğüm sayısı N=11559’dır. Şebekenin düğüm sayısına, R ile aşağıdaki komutları kullanarak hesaplayabiliriz.

>library(igraph)

>DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

>vcount(TB)

[1] 11559

Düğüm sayısı, bilgisayarsız bir ortam da analiz edilemeyecek kadar çok. Düğüm sayısı gerçek-dünya verileri fazla düğüm sayısına sahiptirler..

1. Toplam bağlantısı sayısı L=13750’dir. Şebekenin bağlantı sayısına, R ile aşağıdaki komutları kullanarak hesaplayabiliriz.

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

>ecount(DS)

[1] 13750

Şebekenin bağlantı sayısı düğüm sayısına göre az. Gerçek-dünya şebekelerinde görülen bir durumdur. Bu durum verinin sağlandığı platformun formatından kaynaklanıyor olabilir. Veri, sağlık uzmanı ve ziyaretçilerden oluşmakta. Ziyaretçiler, sağlık uzmanlarına sorular sormaktadır. Bağlantı kısırlığı ziyaretçilerin sorunlarını 1’den fazla sağlık çalışanı ile paylaşmak istemeyişinden olma ihtimali vardır.

Gerçek-dünya şebekelerinde L değeri, Lmax (Lmax: şebekede oluşabilecek bağlantıların tamamı) değerinden çok küçüktür ve bu durum gerçek-dünya şebekelerinin seyrekliğine sebebiyet verir.

1. Düğümlerin ortalama bağlantısı sayısı <k>=2.379099’dir. Şebeke düğümlerinin ortalama bağlantı sayısına, R ile aşağıdaki komutları kullanarak hesaplayabiliriz.

>library(igraph)

>DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

>mean(degree(DS))

[1] 2.379099

Ortalama derece 2.37 çıkması şebekedeki ziyaretçilerin ortalama 2 sağlık uzmanıyla veya sağlık uzmanın ziyaretçilerle bağlantı kurması anlamına gelmektedir.

1. Şebeke düğümlerini özniteliklerini R komutları ile bulalım.

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

>list.vertex.attributes(DS)

[1] “name” “rol” “cinsiyet”

>V(DS)$name #düğümlerin adlarını dökecektir.

>V(DS)$rol #düğümlerin rollerini dökecektir

> V(DS)$cinsiyet #düğümlerin cinsiyetlerini dökecektir.

Sonuçlardan görüleceği gibi “rol” özniteliğinde ‘S’,’V’, “cinsiyet” özniteliğinde ise ‘1’ ve ‘2’ özniteliklerin öznitelikleri olarak atanmıştır.

1. Şebekenin yoğunluğunu R komutlarını kullanarak hesaplayalım.

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

> edge\_density(DS)

[1] 0.00020584

Yoğunluk, düğümler arasındaki karşılıklı bağlantılı olmanın düzeyini belirler. Şebekedeki uyumluluk, birliktelik, dayanışma ve aidiyet şebekenin yoğunluğu ile belirlenir. Şebeke dışardan gelecek tehditlere ve saldırılara ne kadar dayanıklı olduğuna da ulaşılabilir. Bir şebekenin yoğunluğu 0.00020 olarak hesaplanmıştır. Bu oluşabilecek bütün bağlantıların on binde ikisinin gerçekleştiği anlamına gelir. Gerçek-dünya şebekelerinde yoğunluk düşüktür, bu şebeke içinde beklenen bir değerdir.

1. Şebekenin çapını R komutlarını kullanarak hesaplaya biliriz.

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

>diameter(DS)

[1] 19

Şebekenin çapı herhangi 2 düğümün arasındaki en uzak mesafe şebekenin çapını verir. Bu şebeke de 2 düğüm arası en uzak mesafe 19 olarak belirlenmiştir. Bu şebeke göz önünde bulundurulduğunda ziyaretçiler birbirlerini tanımamaları normal. Ancak sağlık uzmanları birbirleri ile bağlantılı olsalardı bu değer normal seviyelere inebilirdi. Stanley Milgram’ın yaptığı deneyde insanlar arası ortalama mesafe altıdır. Günümüzde sosyal ağlar incelendiğinde bu sayı dörtlere kadar düşmüştür.

1. Şebekenin yarıçapını R komutlarını kullanarak hesaplaya biliriz

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

>radius(DS)

[1] 1

Yarıçap; şebeke de düğümler arası en kısa yol. Şebekedeki düğümler arası en kısa mesafe 1 uzaklıktaymış. Bu da demek oluyor ki şebekemizdeki tüm düğümler en az bir düğümle bağlantı kurmuş.

1. Şebeke de yer alan tüm düğümler bağlantılı mıdır? Sorusuna R komutları kullanarak bulabiliriz.

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

>is\_connected(DS)

FALSE

Sonucun FALSE çıkması, şebekedeki tüm düğümler birbirlerine bağlantılı olduğu söylenemez. Gerçek-dünya şebekelerin tüm düğümlerin birbirlerine bağlı değildir. Bizim ağımızda da tüm düğümler bağlantılı değil.

1. Şebekenin bağlantılı bileşen sayısı 14492 adet , bağlantılı bileşenlerin büyüklüğüne ait histogram grafiğini R komutları ile hesaplayıp, çizelim.

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

>components(DS)$no

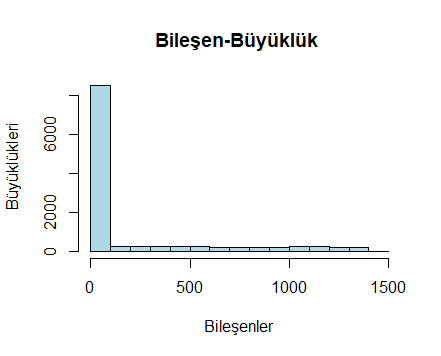
$no #şebekedeki toplam bileşen sayısı

[1] 1406

>hist(components(TB)$membership, col="lightblue",

xlab="bileşenler", ylab=" Büyüklükleri",

main="Bileşen-Büyüklük ") #histogram grafiği çizimi.



Şebekedeki 1406 adet bileşene vardır.. Grafikten de görüldüğü gibi grafikteki ilk sütundaki bileşenler, en yüksek ortalama 8200 düğümlüdür. Diğer bileşenlerin büyüklükleri 2 ile 50 düğüm arasında değişmektedir. Bu durum şebekemizde dev bağlantılı bileşenin varlığına işarettir. Şebekemizde bu durum ziyaretçilerin çoğunluğunun, bazı sağlık uzmanları etrafın da toplandıkları söylenebilir. Bazı sağlık uzmanları dışında kalan sağlık uzmanlarına da az sayıda ziyaretçinin bağlantı kurduğu gözlemlenebilir.

1. Şebeke dev bağlantılı bileşen var mı?

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

>max(components(DS)$csize)

[1] 8209

Bu komutta bileşen boyutlarının en büyüğünü sorguladığımızda sonuç 8209 çıkıyor. Bu da demek oluyor ki şebekedeki bir bileşen 8209 düğümden oluşuyor. Bizim şebekemizde 11559 düğüm var bu da 10’da 7’sine karşılık geliyor. Biz derste şebekenin düğüm sayısının yarısından fazlası bir bileşen de bulunuyorsa o bileşen dev bileşen olabilirdi. Burada şebekenin tüm düğümlerinin yarısından fazlası bir bileşende toplanmış. Bu bileşen dev bağlantılı bileşen diyebiliriz.

1. Şebekenin ortalama patika uzunluğunu R komutları ile hesaplayalım.

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

>mean\_distance(DS, directed = FALSE)

[1] 5.724733

Şebekenin ortalama patika uzunluğu Steve Milgram, Altı adım hipotezine yakın bir sonuç çıkmıştır.. Şebeke de rasgele seçilen 2 kişin arasındaki uzaklık yaklaşık 6 adım olduğu söylenebilir. Çok ilginçtir bu şebeke için de Milgram’ ın araştırma sonuçları tutmaktadır. Buradan da şebekemiz gerçek-dünya şebekelerinin özellikleri ile uyuşmaktadır.

1. Şebekenin yerel kümelenme katsayısı R komutlarını kullanarak hesaplayalım.

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

> mean(transitivity(DS, type = c("local"), isolates = c("zero")))

[1] 0.006972568

Yerel kümelenme katsayısı; şebekeden bir düğümün bağlı olduğu düğümler arasındaki bağlantının varlığından haber verir. Şebekesinin yerel kümelenme katsayısı çok düşüktür. Bu da düğümün komşularının bağlantıları yok denecek kadar az demektir. Bu Şebeke için olağandır, çünkü şebeke ziyaretçiler, sağlık uzmanı ilişkisini vermektedir. Gerçek dünyada da bir doktorun hastaları arasın da bağlantı bulmak zordur. Ama hiç yok değil, bu az olan kısım da hastaların akrabalarına veya arkadaşlarına doktor önermesinden kaynaklı bir bağlantı olabileceği muhtemeldir.

1. Şebekenin derece dağılımını R komutları kulanarak çizelim.

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

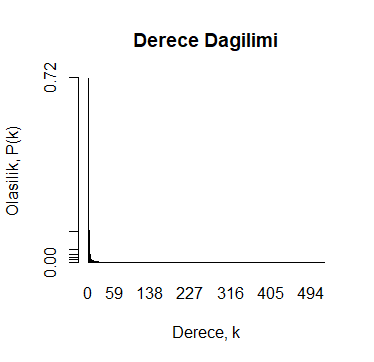
> derd <- degree\_distribution(DS)

>names(derd) <- 0:max(degree(DS))

>barplot(derd, col = "purple",axes = FALSE, xlab = "Derece, k",

ylab = "Olasilik, P(k)", main = "Derece Dagilimi")

>axis(2, at =round(degree\_distribution(DS),2))



Derece dağılımı, ortalama dereceden daha fazla bilgi vermektedir. Bu şebeke için olasılık-derece dağılımı, rasgele seçilen bir düğümün derecesinin k olma olasılığı.

1. Şebeke de düğüm derecesine karşı düğüm yoğunluğu R komutları ile çizelim.

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

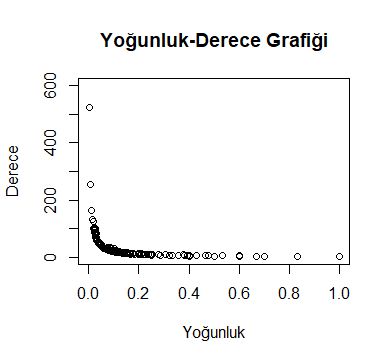
>DS$yogunluk <- sapply(make\_ego\_graph(DS,1),

function (x) {graph.density(x)})

>DS$degree <- degree(DS)

>plot(DS$yogunluk,DS$degree,xlim = c(0,1),ylim = c(0,600),

xlab = "Yoğunluk",ylab = "Derece", main = "Yoğunluk-Derece Grafiği")



Şebekenin merkezi olan düğümü (ve bağlantılarını) dışarıda bırakarak onun bağlantılı olduğu düğümlerin arasındaki şebekenin yoğunluğu olarak hesaplanır (Ders notları,4 Hafta). Grafiğe bakıldığında derecesi en yüksek olan düğümü çıkarıldığında geride kalan düğümlerin bağlantılı olma yoğunluğu sıfıra yakın. Bu durum derece düştükçe yoğunluk artmaktadır.

1. Rasgele 30 tane dersnotlarında 5 hafta önemli notlar var

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

>DS1 <- sample\_degse q(dd, method=c("vl"))

>DS2 <- sample\_degseq(dd, method=c("vl"))

>DS3 <- sample\_degseq(dd, method=c("vl"))

>DS4 <- sample\_degseq(dd, method=c("vl"))

>DS5 <- sample\_degseq(dd, method=c("vl"))

>mean\_distance(DS1, directed = FALSE)

[1] 5.401383

>mean\_distance(DS2, directed = FALSE)

[1] 5.373096

>mean\_distance(DS3, directed = FALSE)

[1] 5.373096

>mean\_distance(DS4, directed = FALSE)

[1] 5.342349

>mean\_distance(DS5, directed = FALSE)

[1] 5.365204

>mean(transitivity(DS1, type = c("local"), isolates = c("zero")))

[1] 0.002123585

>mean(transitivity(DS2, type = c("local"), isolates = c("zero")))

[1] 0.002132585

>mean(transitivity(DS3, type = c("local"), isolates = c("zero")))

[1] 0.002263004

>mean(transitivity(DS4, type = c("local"), isolates = c("zero")))

[1] 0.002441309

>mean(transitivity(DS5, type = c("local"), isolates = c("zero")))

[1] 0.002339735

İlk olarak soruda söylenen 30 adet rasgele şebeke oluşturamadım. For döngüsünü sample\_degseq komutana entegre edemedim. Manuel olarak 5 şebeke oluşturdum. Buradan ortalama patika uzunluklarının yaklaşık olarak DS şebekemizdeki gibi çıktığını görürüz. Oluşturulan rasgele şebekelerin dereceleri eşit olduğundan ortalama patika uzunlukları birbirlerine yakın çıkacaktır. Rasgele ağlar, Gerçek-dünya ağları ile aynı dereceye sahip ise ortalama patika uzunluğunu korurlar.

Yerel kümelenme katsayıları DS şebekesine göre çok düşüktür. Rasgele şebekelerde yerel yapı olmadığından , kümelenme katsayıları çok düşükdür.

1. Şebekenin derece korelasyonunu hesaplayalım.

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

> assortativity\_degree(DS)

[1] -0.09034028

> giant<-decompose(DS)[[1]] #Dev bağlantılı bileşen

> assortativity\_degree(giant) #dev bağ. İçin assortativity

[1] -0.1125436

Bu sonuçtan yüksek dereceye sahip düğümlerin, düşük dereceye sahip düğümler ile bağlantı kurduğunu söyleyebiliriz. Dev bağlantılı bileşen var. Dev bileşen de disassortativity daha küçük ve durum yüksek dereceli düğümler ile düşük dereceli düğümlerin ilişki durumları daha yüksek.

Tüm şebekedeki ziyaretçilerin, müşterisi çok olan sağlık uzmanlarını tercih ettiği sonucuna varabiliriz. Dev bileşende bu durum daha fazla bunun sebebi, dev bağlantı dışındaki ziyaretçilerin, müşteri sayısına bakmadan herhangi bir sağlık uzmanına danışmış olmasından kaynaklanıyor olabilir.

1. Şebekenin Rol ve Cinsiyet korelasyonunu hesaplayalım.

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

> assortativity(DS,types1 = Dugumler$rol)

[1] -0.2711043

> assortativity(DS,types1 = Dugumler$cinsiyet)

[1] -0.375037

> giant<-decompose(DS)[[1]] #Dev bağlantılı bileşen

> assortativity\_degree(giant)

[1] -0.1125436

Roller arası assortativity hesapladığımızda çok düşük bir değer çıkmakta. Bu demek oluyor ki sağlık uzmanları birbirlerine bir şeyler danışmıyorlar. Aynı şekilde ziyaretçilerinde birbirlerine bir şeyler danışmamaktadırlar. Bu durum iyi olabilir ama sağlık uzmanlarının kendi aralarında bilgi paylaşımının olmaması büyük problem. Sonuç; genel de ziyaretçiler sağlık uzmanları ile bağlantıya geçmiştir.

Cinsiyeteler arası assortativity hesaplandığında rollere göre daha düşük çıkmaktadır. Bu da demek oluyorlar ki ziyaretçiler danışmak istediği sağlık uzmanının cinsiyetine bakmamakta.

1. Şebekemizi Rol ve Cinsiyet çizgelerini R komutları ile çizelim.

>library(igraph)

> DS<-graph\_from\_data\_frame(Baglantiler,directed = FALSE,vertices = Dugumler)

>par(mfrow=c(1,1),mar=c(0.1,0.1,0.1,0.1))

##Cinsiyet için çizge

>V(DS)$shape<-"circle"

>V(DS)[cinsiyet==1]$color<-"pink1" #bayan pembe

>V(DS)[cinsiyet==2]$color<-"skyblue1" #erkek mavi

>DS1<-V(DS)[cinsiyet==1]

>DS2<-V(DS)[cinsiyet==2]

>E(DS)[DS1 %--% DS1]$color<-"chartreuse" #bayan-bayan fst.yeşil

>E(DS)[DS2%--% DS2]$color<-"darkorchid3" #erkek-erkek mor

>E(DS)[DS1 %--% DS2]$color<-"red1" #bayan-erkek kırmızı

>l<- layout\_with\_fr(DS,grid = c("nogrid"))

>plot(DS,vertex.label ="", vertex.size=1, edge.width =1, edge.arrow.size = 1,layout=l )

##Rol için çizge

>V(DS)$shape<-"circle"

>V(DS)[rol==S]$color<-"gray75" #sağlı uzamanı gray

>V(DS)[rol==V]$color<-"khaki2" #ziyaretçi halki

>DSr1<-V(DS)[rol=="S"]

>DSr2<-V(DS)[rol=="V"]

>E(DS)[DSr1 %--% DSr1]$color<-"pink" #s.uzm-s.uzm penbe

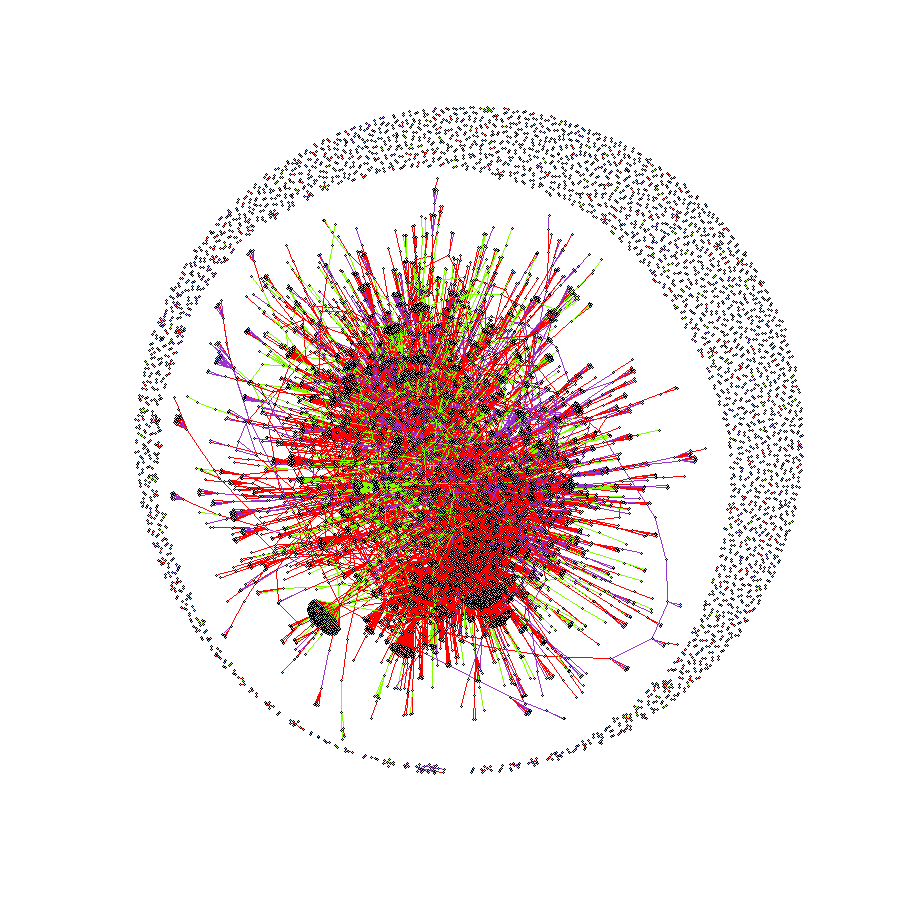
>E(DS)[DSr2%--% DSr2]$color<-"lightblue" #ziyare.-ziyare. mavi

>E(DS)[DSr1 %--% DSr2]$color<-"yellow" #s.uzm-ziyare. sarı

>plot(DS,vertex.label ="", vertex.size=1, edge.width =1,

edge.arrow.size = 1,layout = l)

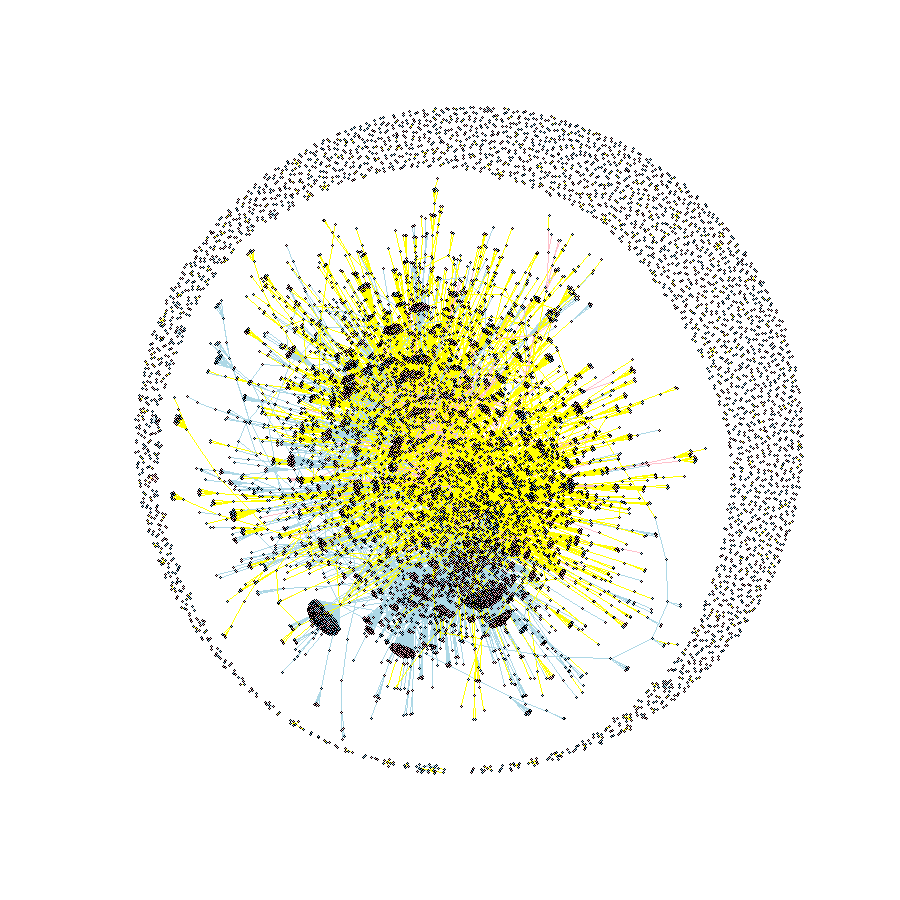
* 1. Cinsiyet için çizgemiz.



17. soruda derece korelasyonuna bakmıştık ve disassortativity çıkmıştı. Çizgede de düğümlerin çoğunluğunun derecesi yüksek olan düğümlerle bağlantı kurduğu görülür. Çizgende görülüyor ki düğümlerin derece bakımından disassortativity mevcuttur. Şebekemiz dev bağlantılı bir şebekedir. Çizgede görüldüğü gibi düğümler çizgenin orta kısmında yığılmışlardır. Dev bileşende de derece disassortativity vardır. Çoğunluk derecesi yüksek olan düğümlerle bağlantı kurmuşlar.

18. soruda cinsiyet göre disassortativity olduğunu söylemiştim. Çizgede kırmızı bağlantı rengi kadın ve erkekleri temsil etmekte. Çizgedeki bağlantılar da kırmızının yoğunlukta olduğu görülür. Bu çizge cinsiyet çizgesi olduğundan roller için yorum yapmak zor.

* 1. Rol için çizgemiz.



17.soruda derece korelasyonuna bakmıştık aynı şekilde bu çizgeden de görülüyor ki düğümler derecesi yüksek düğümlerle bağlantı kurmuştur.17.sorudaki gibi disassortativity geçerliliğini sürdürmektedir. Dev bağlantılı bileşenimiz içinde geçerlidir.

18.soruda roller arası korelasyon negatif çıkmıştı. Rol çizgemizde ziyaretçi-sağlık çalışanı arasındaki bağlantıyı sarı renk temsil etmektedir. Çizgemizde görüldüğü üzere bağlantılarda sarı rengin baskın olduğu görülmektedir. Burada da disassortativity geçerliliği devam etmekte. Bu durum çok normal, genelde ziyaretçiler sisteme sağlık uzmanlarına danışmak için giriş yaptığı için roller arası disassortativity çıkması olağandır.

# TEŞEKKÜRLER…