

Redes não orientadas - Gestão de contactos sociais diretos entre habitantes de uma zona residencial

Trabalho de Grupo I realizado no âmbito da Unidade Curricular
de Análise de Redes do 3º ano da Licenciatura em Ciência de
Dados

Allan Kardec Rodrigues, 103380
aksrs@iscte-iul.pt

André Plancha, 105289
Andre_Plancha@iscte-iul.pt

Diogo Freitas, 104841
daafs@iscte-iul.pt

João Francisco Botas, 104782
Joao_Botas@iscte-iul.pt

Marco Esperança, 110451
mdeao@iscte-iul.pt

Índice

Introdução 2

Q1: Análise da rede 3

Q2: Análise da componente gigante 4

Q3 4

Anexos 4

Introdução

O presente trabalho pretende analisar uma rede não orientada, que representa os contactos sociais diretos entre os habitantes de uma zona residencial. Os nodos (N) representam os habitantes da zona residencial e cada ligação (L) representa um contacto social direto entre dois habitantes. Na primeira questão pretende-se a análise de toda a rede e na segunda questão a componente gigante. Para isto, foi utilizado o *package* *igraph* da linguagem de programação R. Para importar esta rede para o R, utilizou-se a função `read_graph` da seguinte forma:

```
library(igraph) # importação da biblioteca
rede <- read_graph("trab_links.txt", format = c("edgelist"), directed=F)
plot(rede, vertex.size = 7, vertex.label.cex = .35) # com estes args para reduzir tamanho do grafo
```

A rede é representada por:

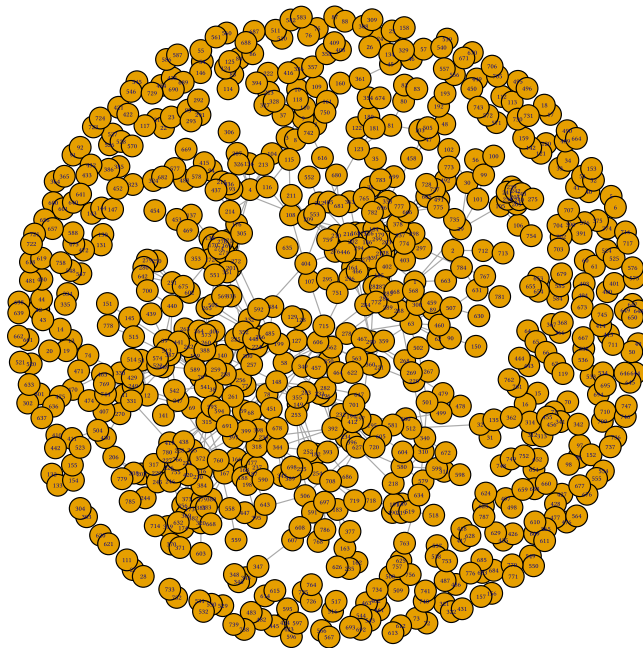


Figura 1: Representação visual da rede

A análise de uma rede pode ser útil para:

- **Identificação de Grupos e Comunidades:** Pelas componentes conectadas é possível descobrir que grupos de pessoas interagem mais frequentemente entre si, indicando possíveis comunidades na zona residencial.
- **Centralidade dos Indivíduos:** Através de algumas métricas de centralidade como o grau (número de conexões) ou a centralidade de informação (quão importante é um indivíduo no fluxo de informações) podem ajudar a identificar pessoas mais influentes ou centrais na rede social.
- **Propagação de Informação ou Ideia:** Estudar como uma informação ou ideia pode se espalhar na comunidade através dos contatos sociais diretos.
- **Identificação de Pontos Críticos:** Encontrar os nós mais importantes na rede cuja remoção poderia interromper a propagação de informação ou de uma influência específica.
- **Monitorizar Mudanças:** Observar como a rede evolui ao longo do tempo, para oferecer *insights* sobre mudanças nas relações sociais e na estrutura da comunidade.

A segunda rede é a **componente gigante**, isto é, a maior componente conexa da rede original. Uma componente conexa numa rede é um conjunto de vértices onde cada par de vértices está conectado por um caminho. Assim, na componente gigante, cada vértice está acessível a partir de qualquer outro vértice dentro dessa componente através de um caminho.

A identificação da componente gigante é essencial para compreender a estrutura e conectividade de uma rede, destacando a sua parte mais extensa e predominante, revelando a sua coesão estrutural. Assim, no contexto deste trabalho, pode ser muito útil para representar uma comunidade social coesa, permitindo detetar os contactos sociais mais frequentes e fortes, permitindo evidenciar uma forte interação entre os habitantes.

A representação da componente gigante encontra-se de seguida:

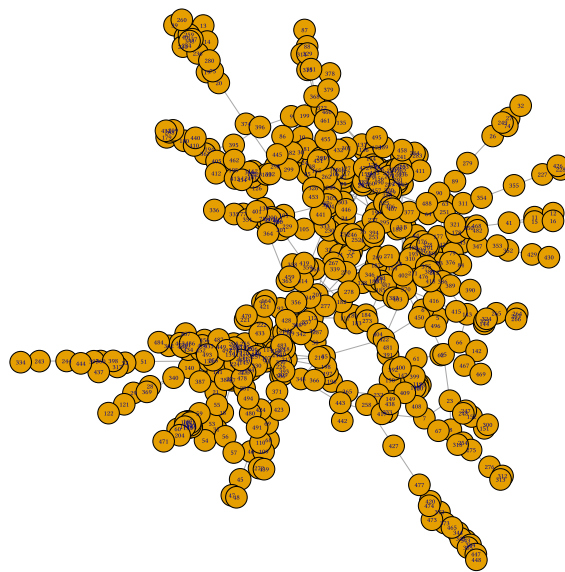


Figura 2: Representação visual da componente gigante

Q1: Análise da rede

Na questão 1, primeiramente, pretende-se indicar a dimensão, e o número de ligações, assim como a densidade, o grau médio e a distribuição do grau.

O grau de um nodo i de uma rede não orientada é o número de ligações incidentes nesse nodo e representa-se por k_i . O grau médio é a média dos graus dos nodos de uma rede e representa-se por $\langle k \rangle$.

A densidade é uma medida relativa que relaciona o número de ligações existentes numa rede com o número máximo possível de ligações. Com a rede é não orientada, esta pode ser calculada por $d = \frac{2L}{N(N-1)}$.

Pela inspeção visual da representação gráfica da rede apresentada na Introdução pode-se constatar que é uma rede de dimensões consideráveis e não se conseguindo identificar os caminhos a olho nu. Isto pode ser corroborado pelo número de vértices da rede que são 787 e o número de arestas que são 1197.

Relativamente ao grau médio, $\langle k \rangle$, é de 3.041931, o que significa que, em média, cada nodo está ligado a aproximadamente 3 nodos. No contexto do problema, significa que cada habitante na zona residencial está diretamente conectado a cerca de 3 outros habitantes através de contactos sociais diretos, representados na rede.

Quanto à densidade obteve-se o valor de 0.003870142, um valor muito inferior a 1, pelo que a rede é esparsa. Desta forma, existe apenas uma pequena fração do número total de conexões possível entre os habitantes que está realmente representada na rede. Além disso, pode indicar também a presença de subgrupos ou comunidades dentro da sua zona residencial que têm poucos contactos fora dos seus grupos, resultando nessa densidade reduzida na rede global.

Q2: Análise da componente gigante

Q3

Anexos