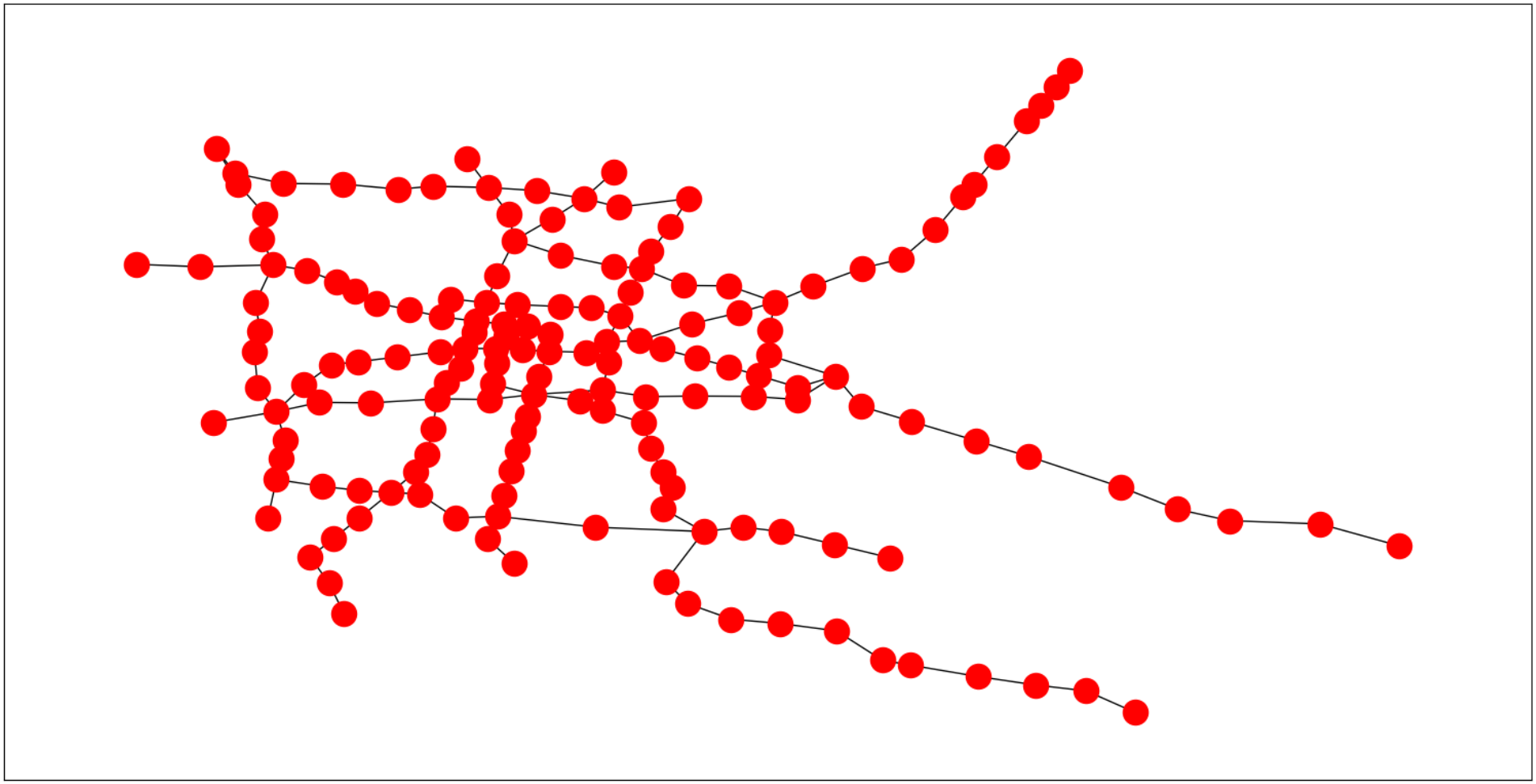
Resiliencia y robustez en redes

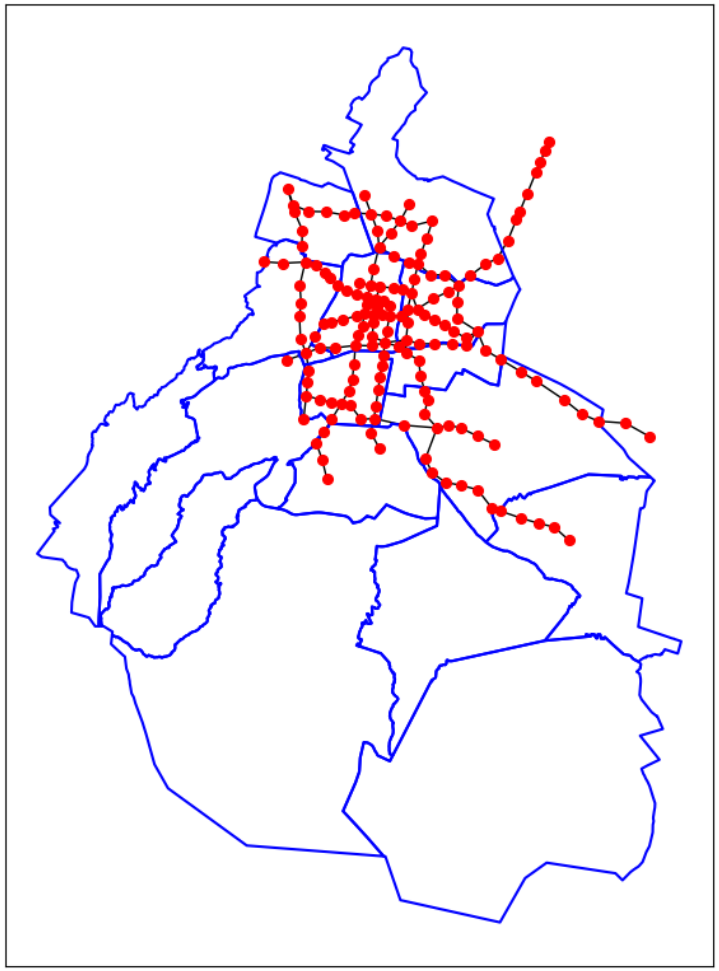
Módulo 4 : Técnicas computacionales avanzadas para modelar fenómenos sociales
Concentración en Economía Aplicada y Ciencia de Datos
ITESM

13 de mayo de 2023

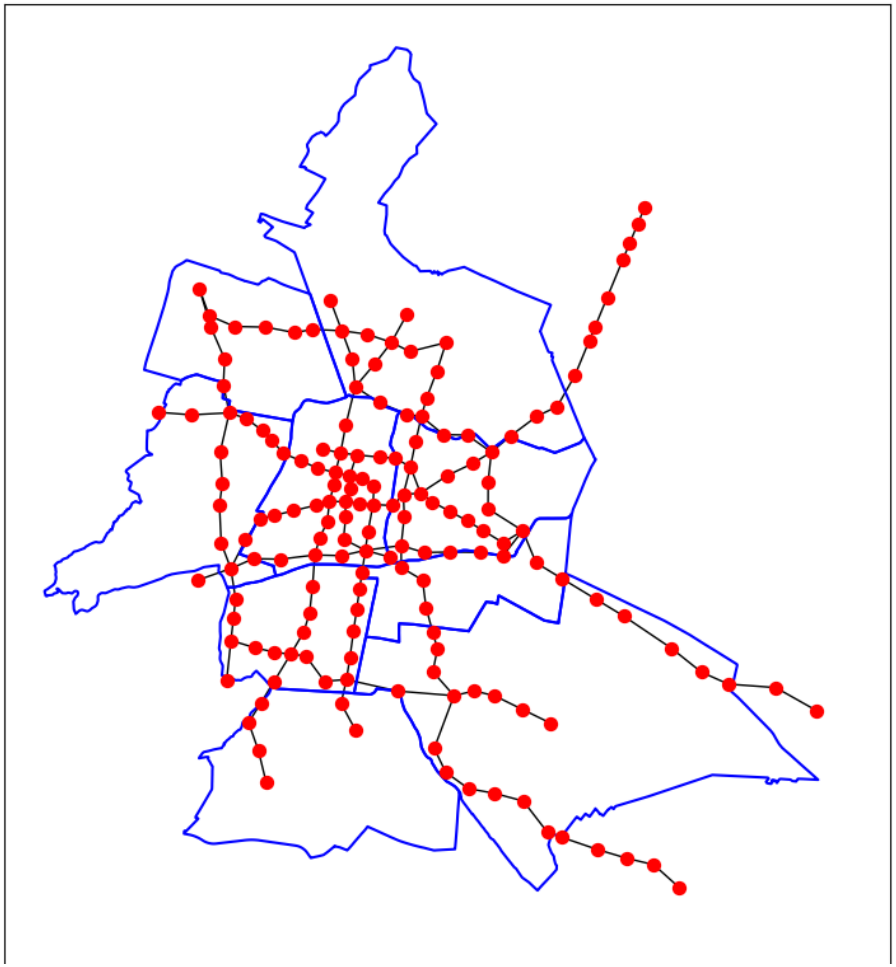
¿Qué red es?



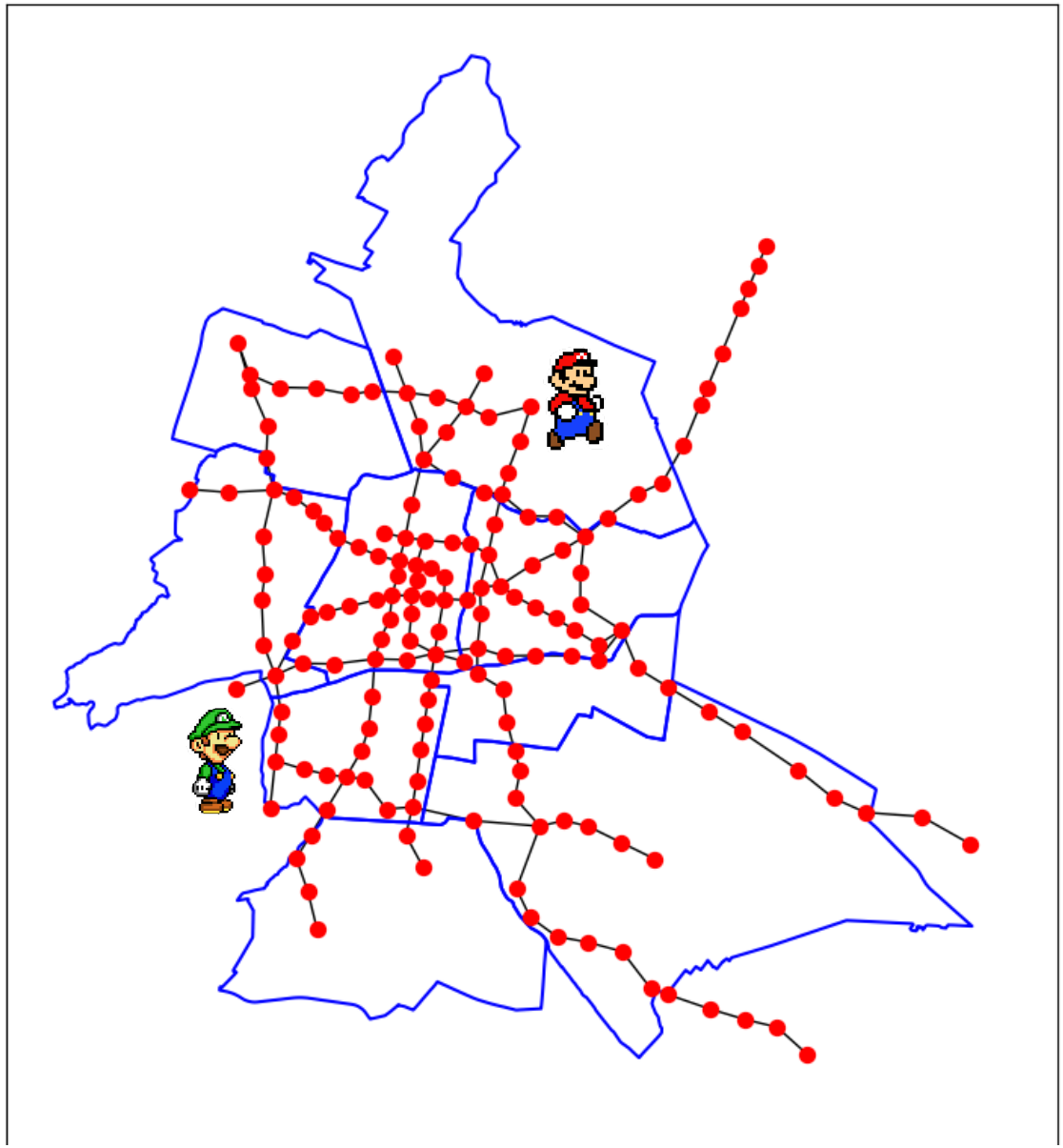
¿Ahora?



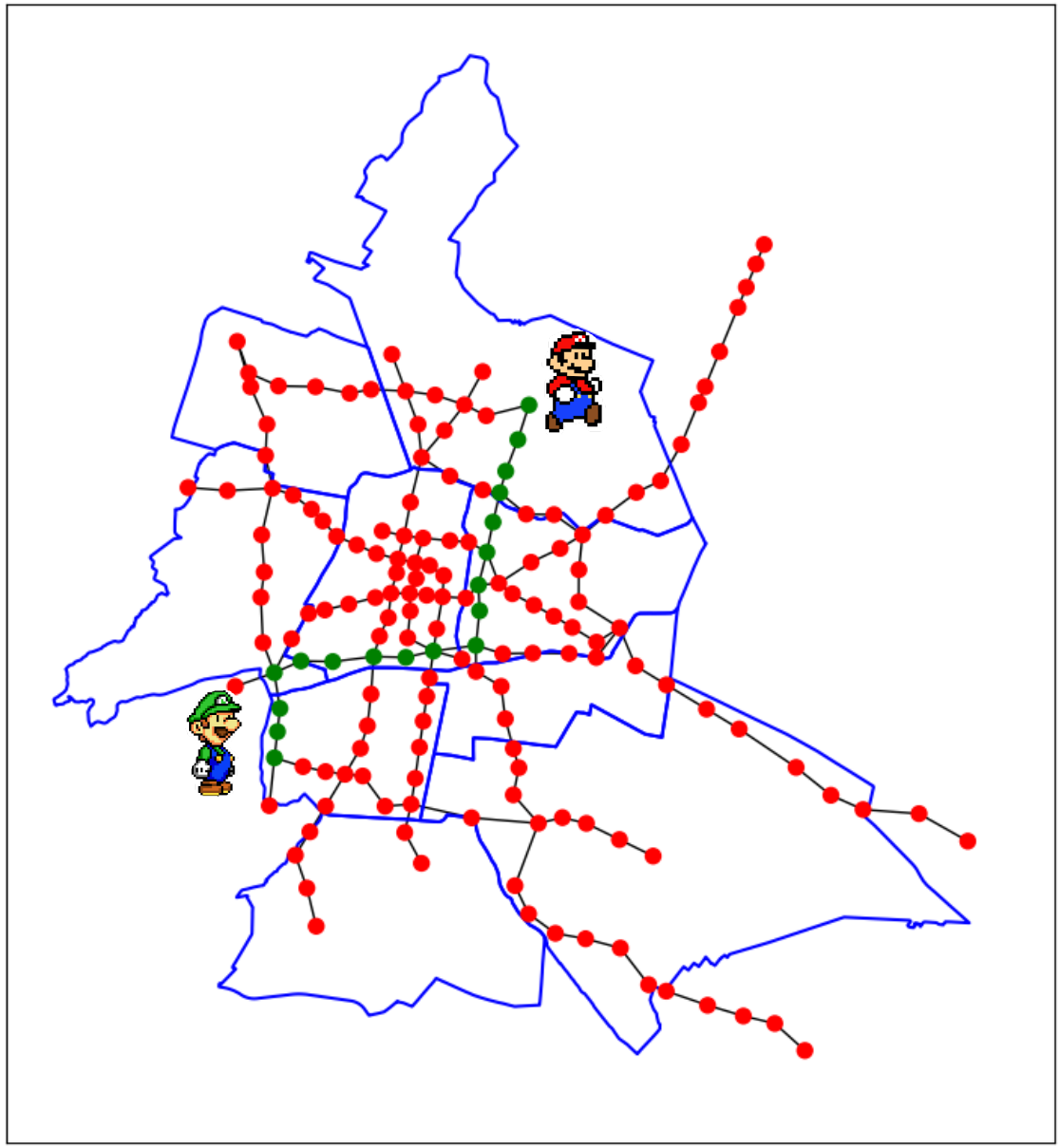
¿Ahora?



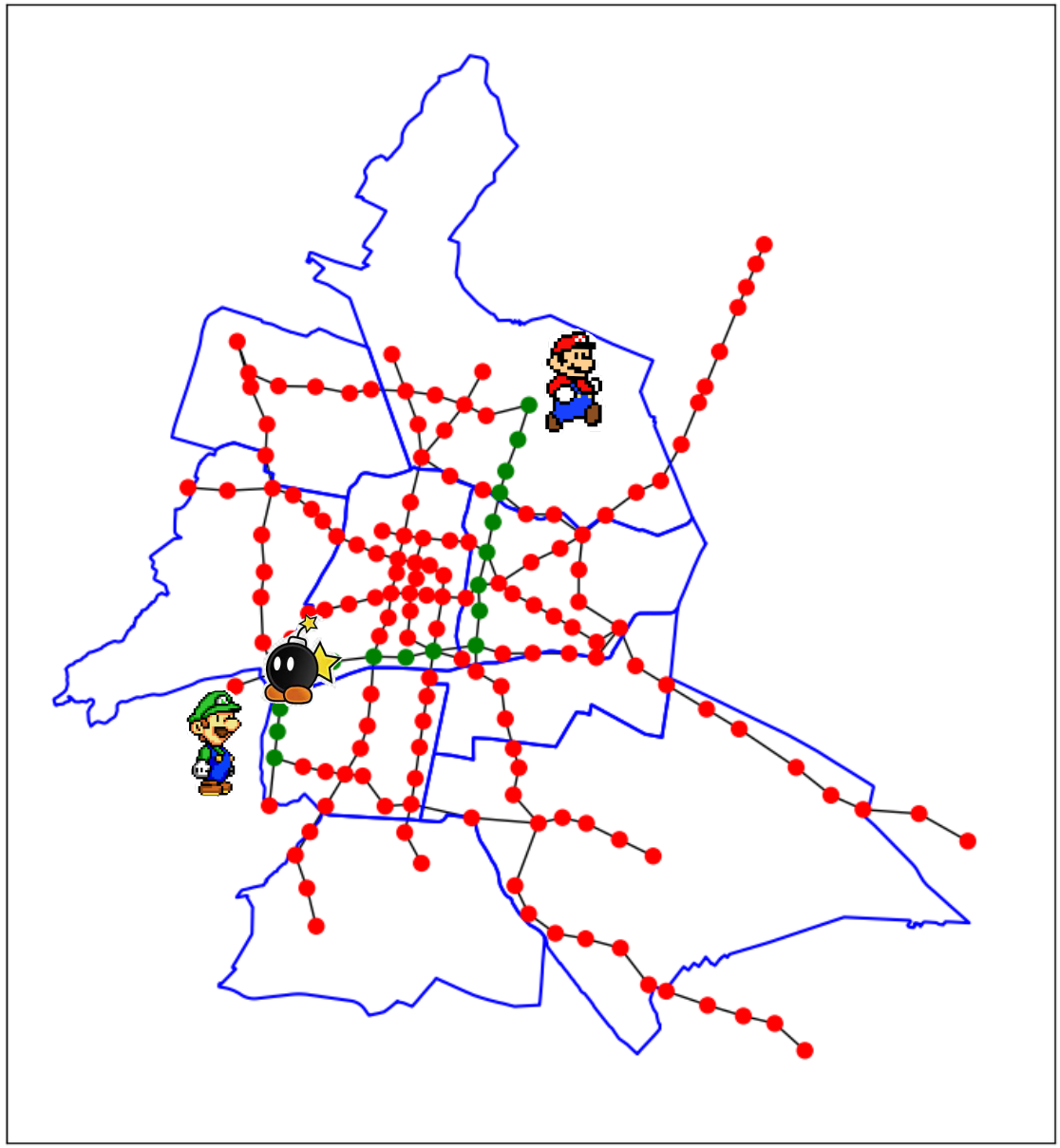
Luigi quiere visitar a Mario



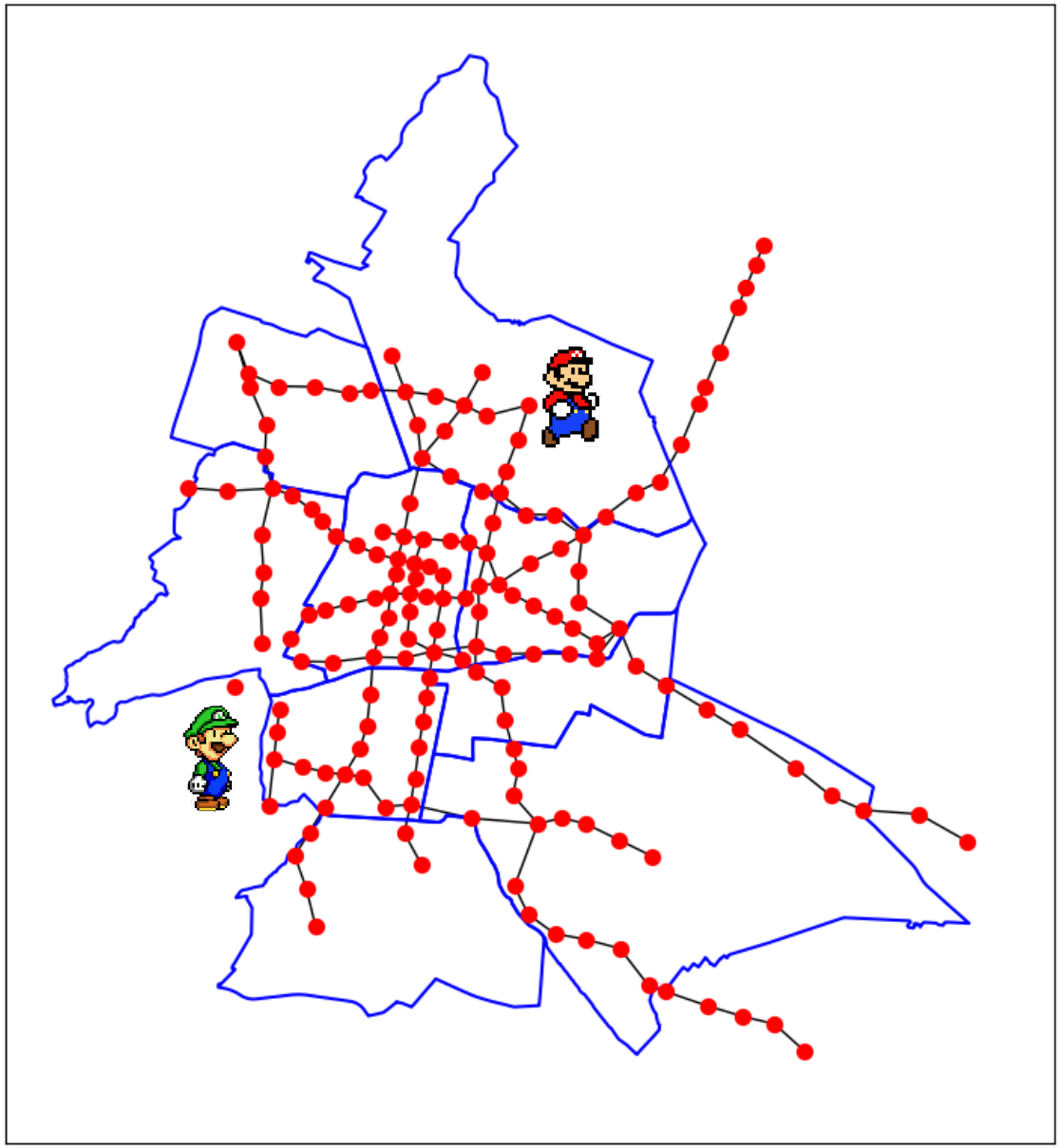
Luigi quiere visitar a Mario



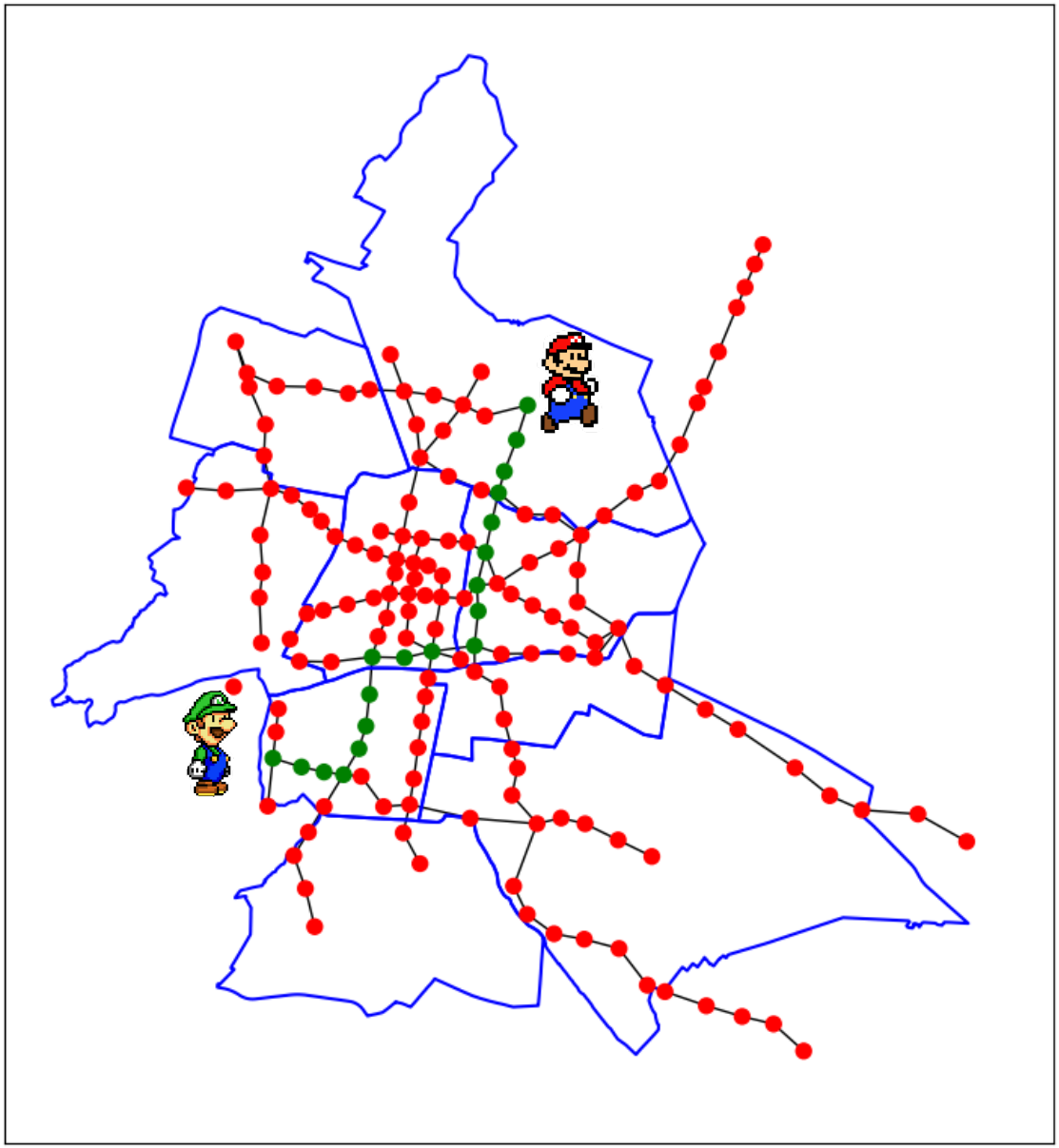
Luigi quiere visitar a Mario



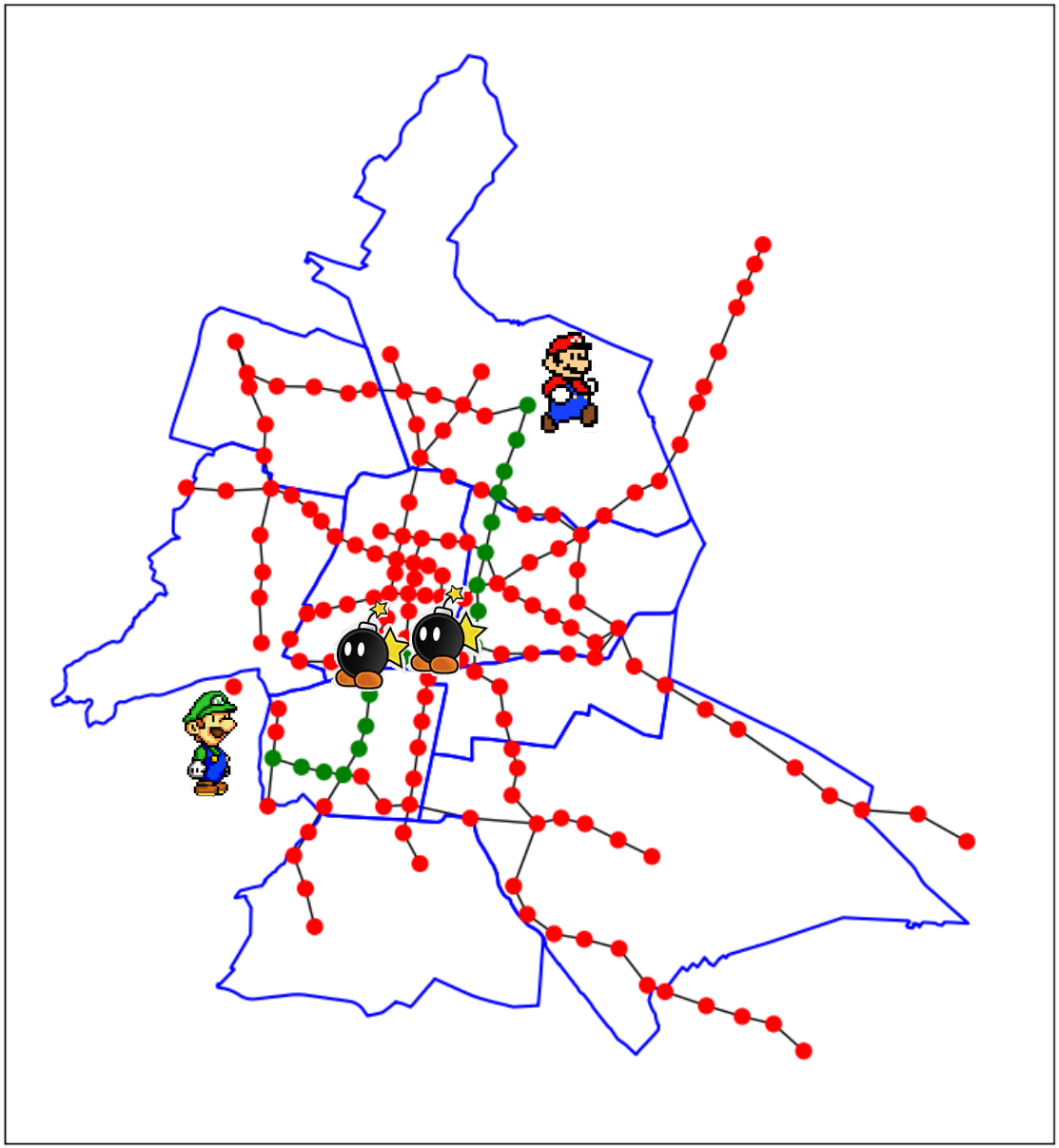
Luigi quiere visitar a Mario



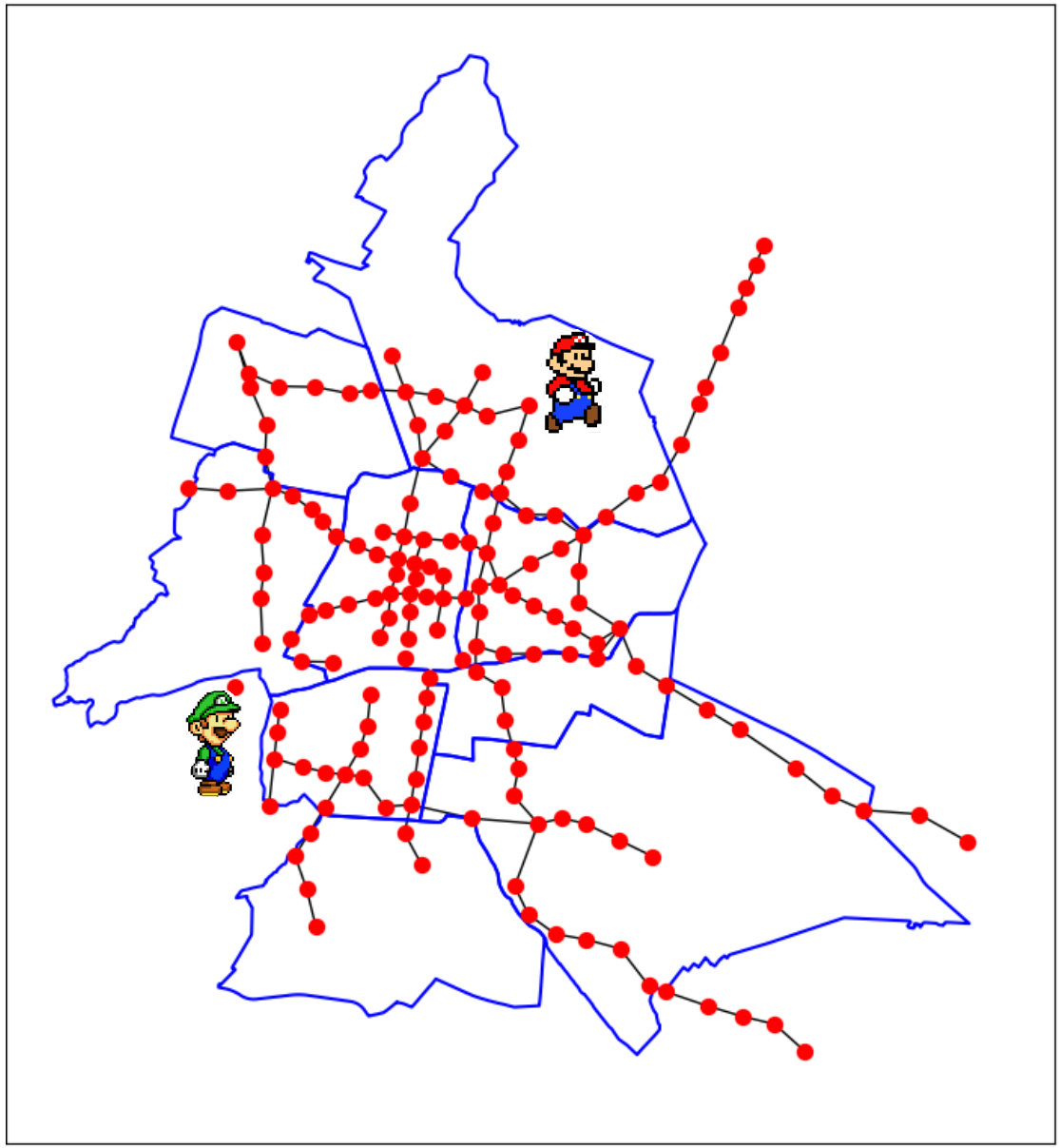
Luigi quiere visitar a Mario



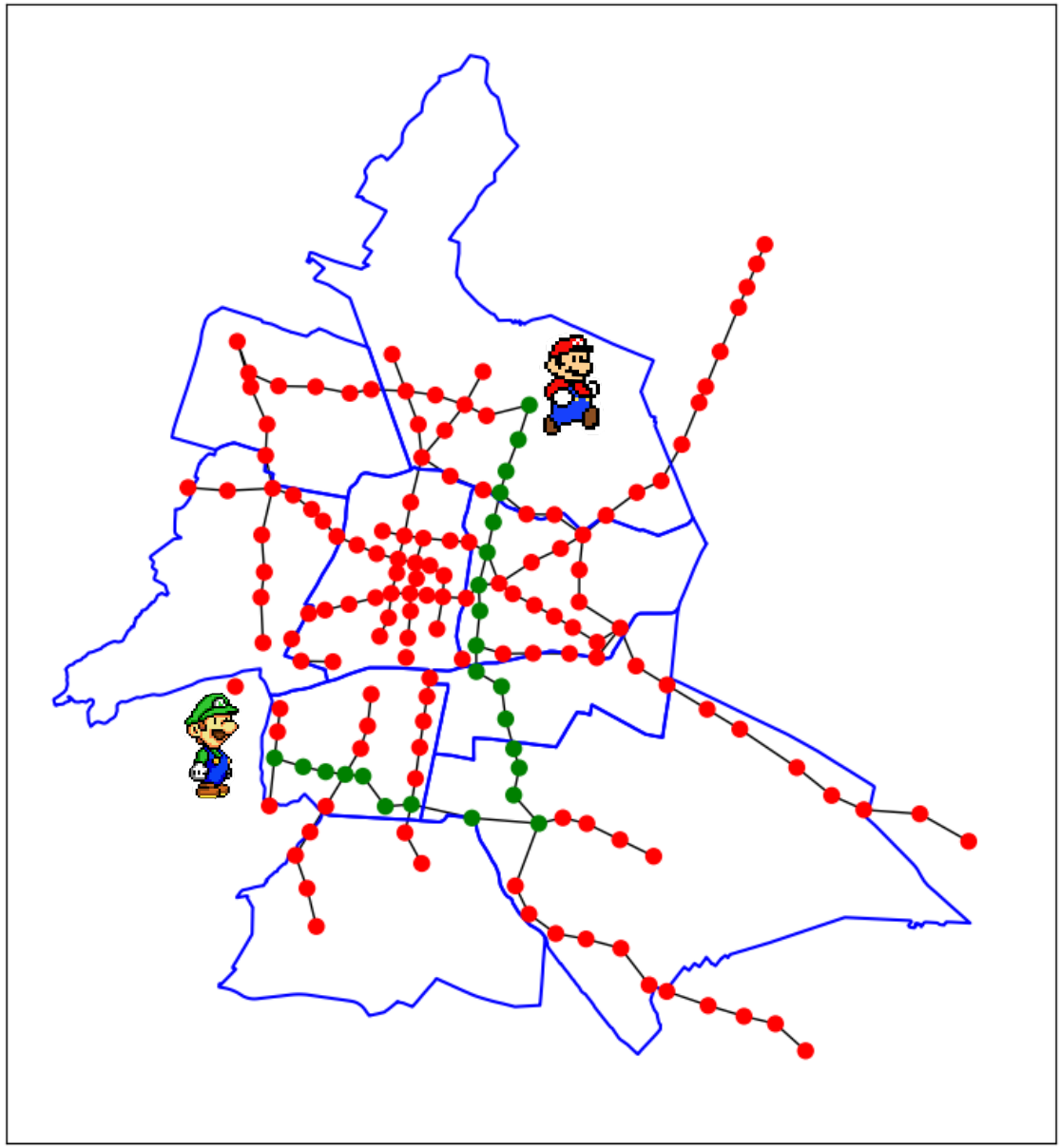
Luigi quiere visitar a Mario



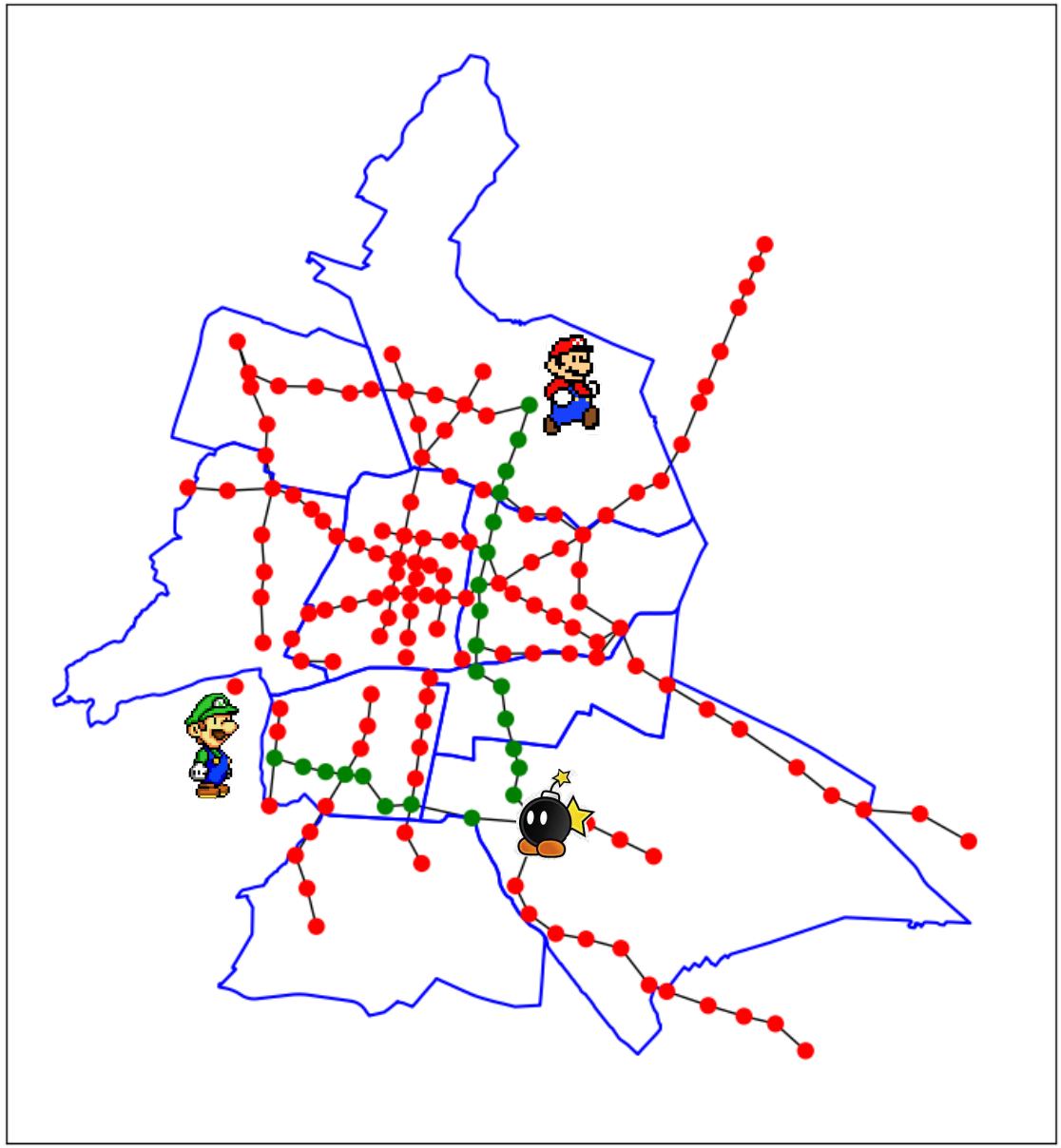
Luigi quiere visitar a Mario



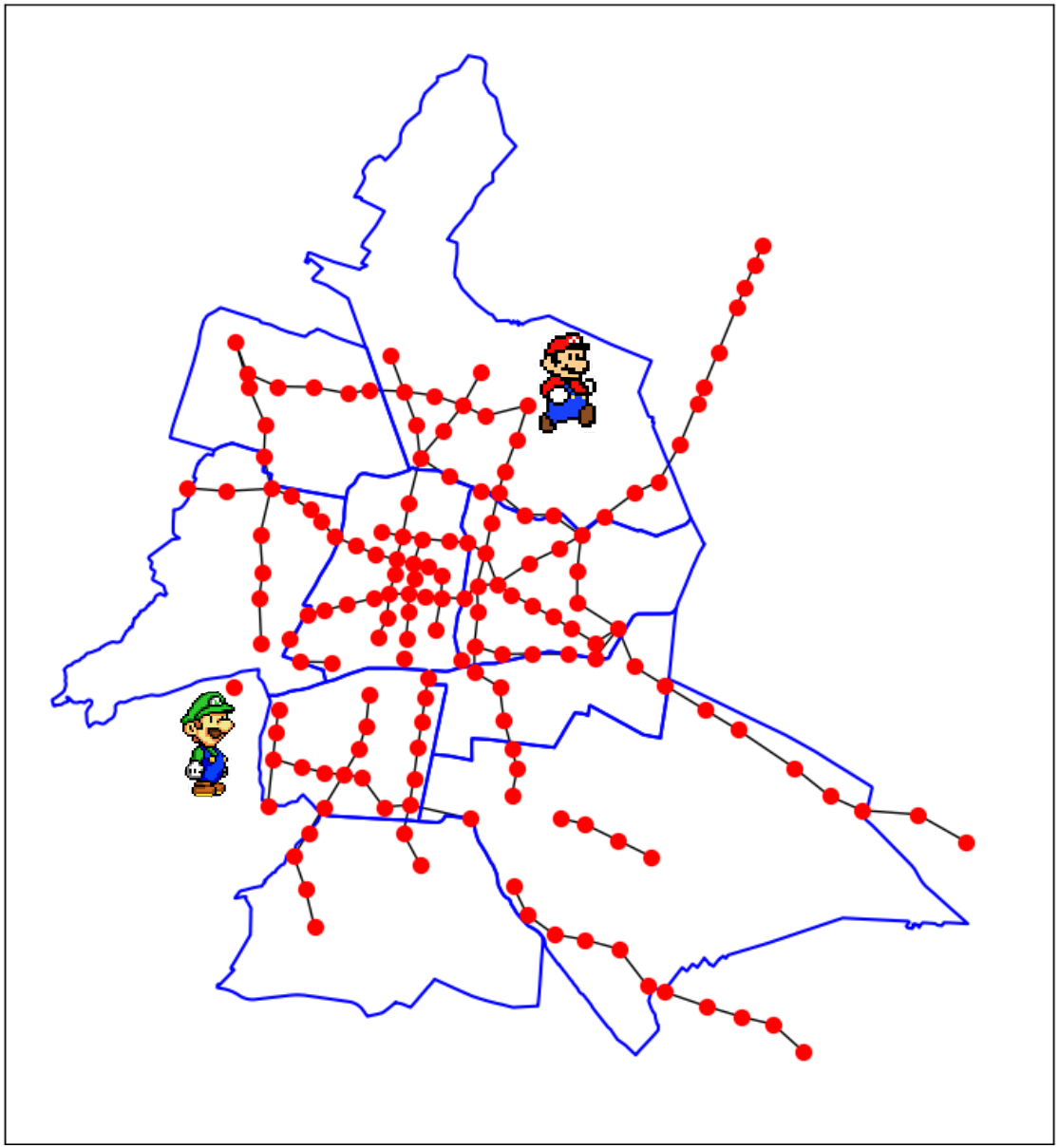
Luigi quiere visitar a Mario



Luigi quiere visitar a Mario



Luigi quiere visitar a Mario



Luigi quiere visitar a Mario

Resiliencia y robustez en redes

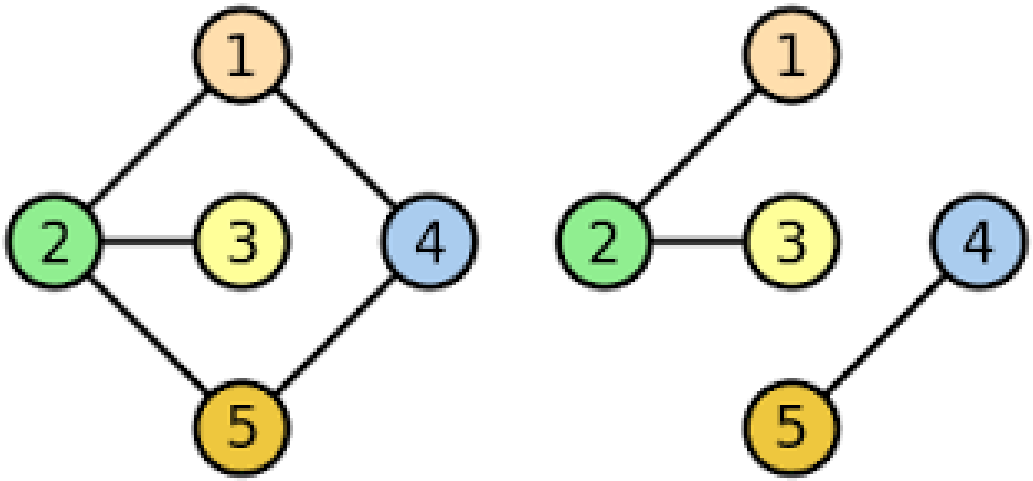
Un sistema es **robusto** si la falla de alguno de sus componentes no afecta su funcionamiento (Menczer et al., 2020).

Resiliencia y robustez en redes

- ¿Cómo definimos la robustez de una red?
- Asumiendo que alguno de los nodos deja de funcionar, podemos preguntarnos cómo cambia la estructura y el funcionamiento de la red sin ese nodo y sus conexiones.
- Una primera evaluación corresponde a estudiar el efecto de remover nodos y observar el cambio en las propiedades topológicas de la red (Menczer et al., 2020).

Resiliencia y robustez en redes

- Una propiedad topológica de las redes es la **conexidad** de la red.
- Una red es **conexa** si todos los nodos son alcanzables desde cualquier nodo de la red.
- Si una red es no conexa, esta tiene dos o más componentes conexos.



Resiliencia y robustez en redes

- El internet es una red conexa. En caso contrario, sería imposible enviar paquetes entre routers.
- De esta manera, una forma de definir y medir la robustez de una red es observar cómo la eliminación de un nodo y sus conexiones afecta la conexidad del sistema.

Resiliencia y robustez en redes

- Si el sistema permanece conexo, podemos asumir que este se mantendrá funcionando correctamente, hasta cierto punto.
- Sin embargo, una ruptura de la red en componentes desconectados indicaría un severo daño en la red que podría comprometer su funcionamiento.
- La prueba estandar de robustez para redes consiste en verificar como es afectada la conexidad de la red en la medida que más y más nodos son removidos.

Componente gigante

Para estimar la cantidad de daño ocasionada, se calcula el tamaño relativo del **componente gigante** (Barrat et al., 2008), (*i.e* la razón entre la cantidad de nodos en el componente gigante, S_f, y la cantidad de nodos inicialmente en la red, $S_0 = N$).[1]

$$\frac{S_f}{S_0} \tag{1}$$

La red mantendrá su capacidad para realizar su tarea mientras el tamaño del componente gigante sea igual al tamaño inicial de la red, $\frac{S_f}{S_0} = 1$.

[1] El **componente gigante** de una red es el componente conexo más grande de la red.

Componente gigante

Cuando $S_f < S_0$ la red habrá sido rota en componentes pequeños disconexos y no se garantiza el funcionamiento de la red.

El estudio de la evolución de la razón $\frac{S_f}{S_0}$ como una función de la fracción de nodos removidos f caracteriza la respuesta de la red al daño.

Error and attack tolerance of complex networks (Albert et al., 2000)

- Albert et al. (2000) analizaron el efecto de ataques aleatorios y dirigidos en redes homogéneas y heterogeneas.

Redes Homogéneas

- Las redes exponenciales son redes homogéneas en el sentido que se caracterizan en que la mayoría de sus nodos tienen la misma cantidad de conexiones.
- El modelo Erdős–Rényi (Barabasi and Albert, 1999)erdos59a puede construir este tipo de redes.

Redes Homogéneas

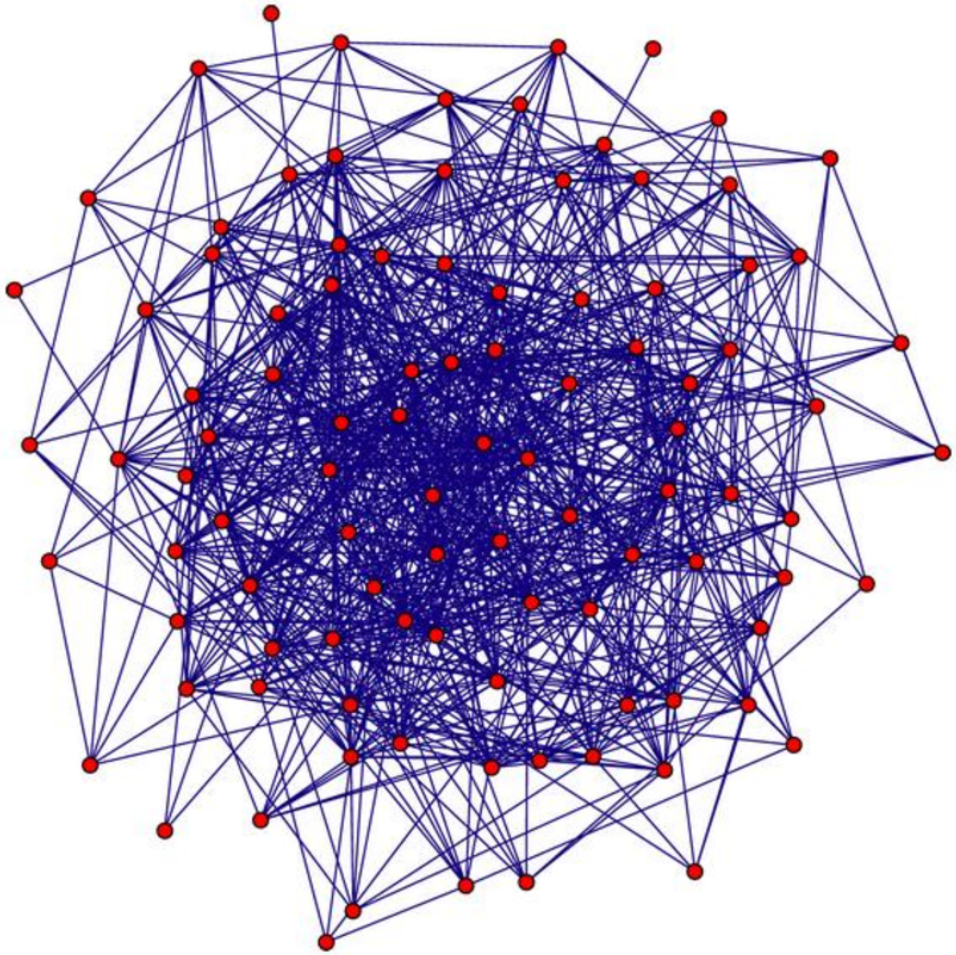


Figura: Red Homogénea

Redes Heterogéneas

- Las redes libres de escala son heterogéneas en el sentido que la mayoría de sus nodos tienen pocas conexiones mientras que pocos nodos tienen muchas conexiones.
- El modelo Barabasi-Albert (Barabasi and Albert, 1999) construye redes libres de escala.

Redes Heterogéneas

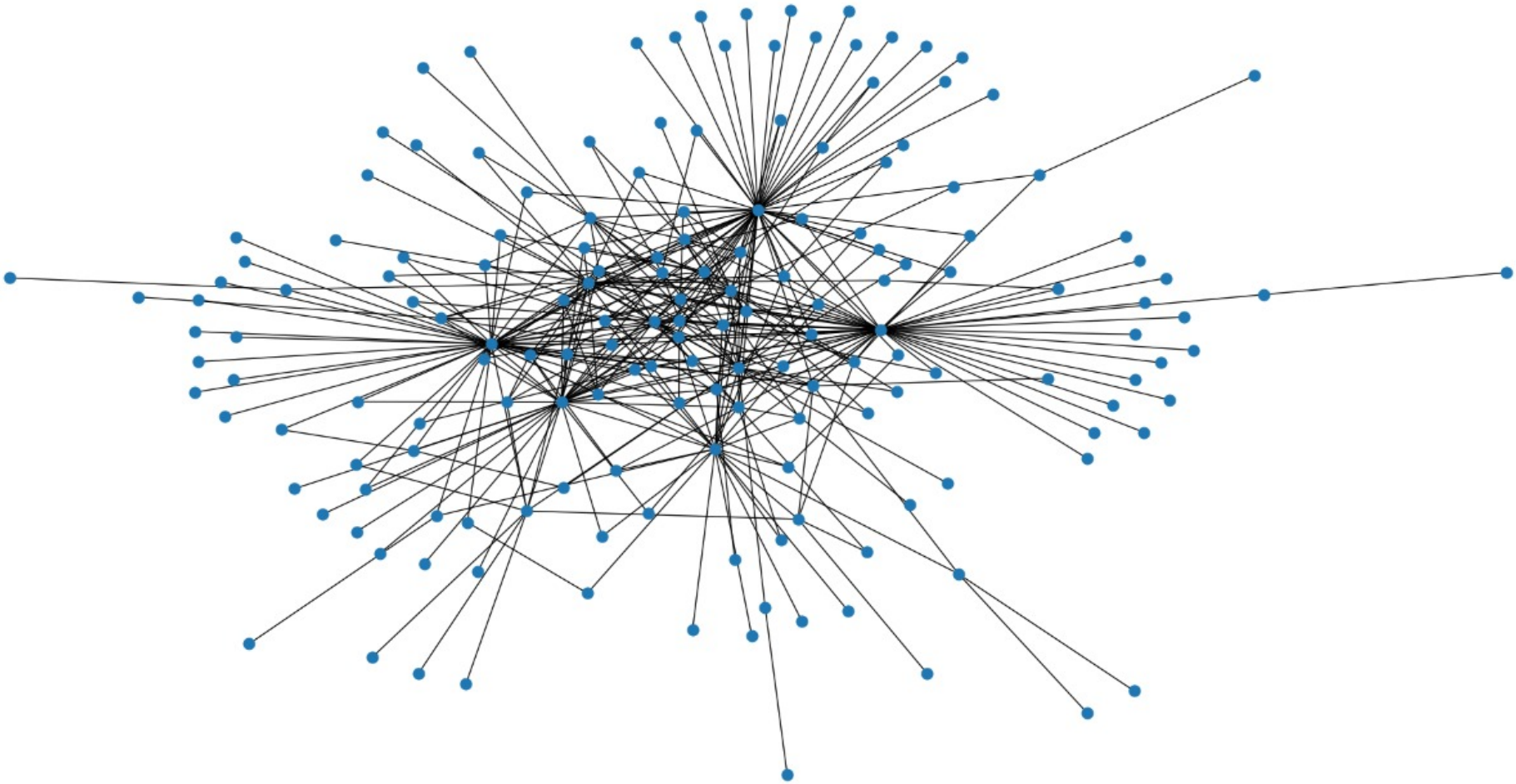


Figura: Red libre de escala (Heterogénea)

Redes libres de escala

- Las redes que presentan una distribución de grado que sigue una ley de potencia se les conoce como *redes libres de escala* (Estrada and Knight, 2015).
- Para el caso de variables continuas, la distribución de Pareto sigue una ley de potencia (Masuda and Lambiotte, 2020):

$$p(x) = Cx^{-\alpha} \qquad (x \geq x_{min}) \tag{2}$$

donde α es el exponente de la ley de potencia de la distribución, $C = (\alpha - 1)x_{min}^{\alpha-1}$ es la constante de normalización.

- Otra propiedad de las distribuciones que siguen una ley de potencia es que no pueden ser caracterizadas con sus momentos (Masuda and Lambiotte, 2020).

Redes libres de escala

Una característica más de las distribuciones que siguen una ley de potencia es que en un plano log – log toman la forma de una línea recta porque al aplicar logaritmo a la ecuación (1) obtenemos:

$$\log p(x) = \log C - \alpha x \tag{3}$$

Cuando se quiere probar que datos empíricos están distribuidos como una ley de potencia, se suele graficar su distribución en la escala log – log y estimar el valor de α.

Modelo Barabasi-Albert (1999)

El modelo Barabási-Albert (BA) es conocido por generar redes libres de escala en la que la probabilidad que un nodo tenga k aristas sigue una una distribución de ley de potencia del tipo $p_k \sim k^{-\gamma}$, con exponente $\gamma = 3$.

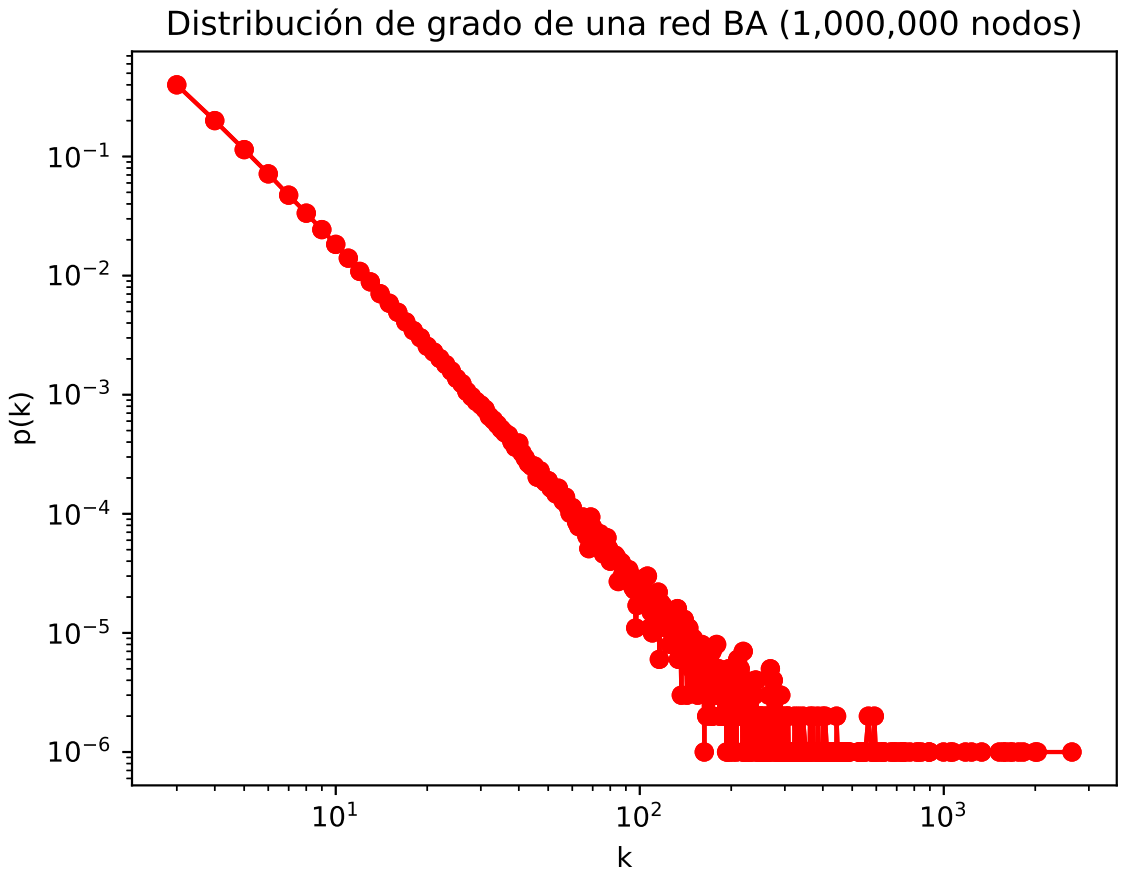


Figura: Distribución de grado en una red BA

Modelo Barabasi-Albert (1999)

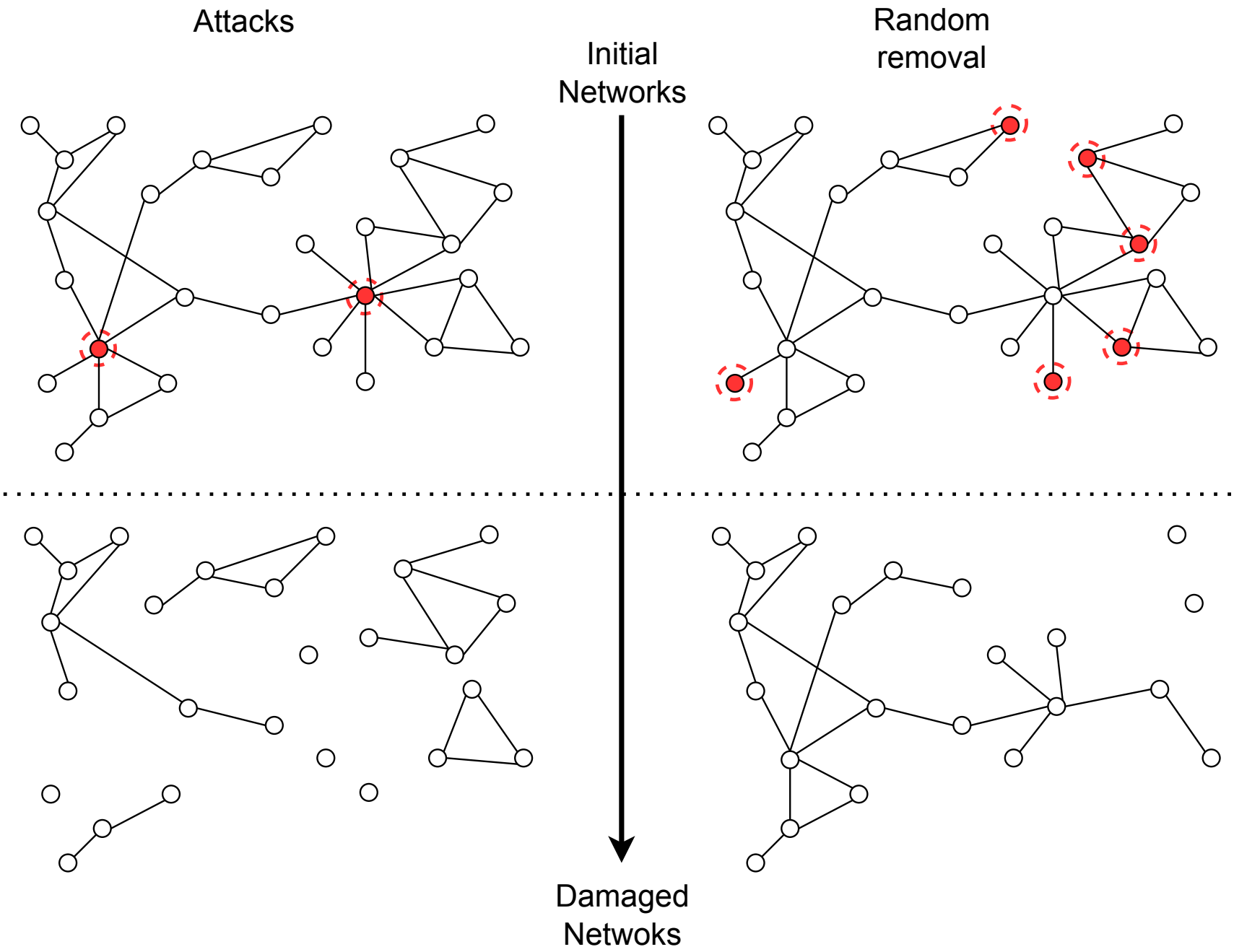
- El modelo BA consiste en crecer una red comenzando con una cantidad pequeña de nodos m_0, y en cada paso de tiempo se agregan nuevos nodos con $m(\leq m_0)$ aristas que se vinculan a los nuevos nodos.
- Los autores asumen que la probabilidad que un nuevo nodo sea conectado a un nodo i dependende la conectividad (grado) de ese nodo k_i (*preferential attachment*), de manera que $\Pi(k_i) = \frac{k_i}{\sum_j k_j}$.

Simulación

https://sarah37.github.io/barabasialbert/

Error and attack tolerance of complex networks (Albert et al., 2000)

- Albert et al. (2000) consideraron el daño ocasionado por remover cierta cantidad de nodos o conexiones, ya sea de forma aleatoria (para modelar fallas aleatorias) o por ataques dirigidos (para modelar daños intencionales).
- Los ataques dirigidos son realizados en función de la **centralidad** (ya sea medida por el grado del nodo o por la centralidad betweenness) y puede describir ataques intencionales que tienen como objetivo maximizar el daño en la red al enfocar el ataque hubs importantes.
- Albert et al. (2000) encontraron que las redes homogéneas y heterogéneas reaccionan diferente al incrementar los niveles de daño.


Attacks
Initial
Networks
Random
removal
Damaged
Netwoks

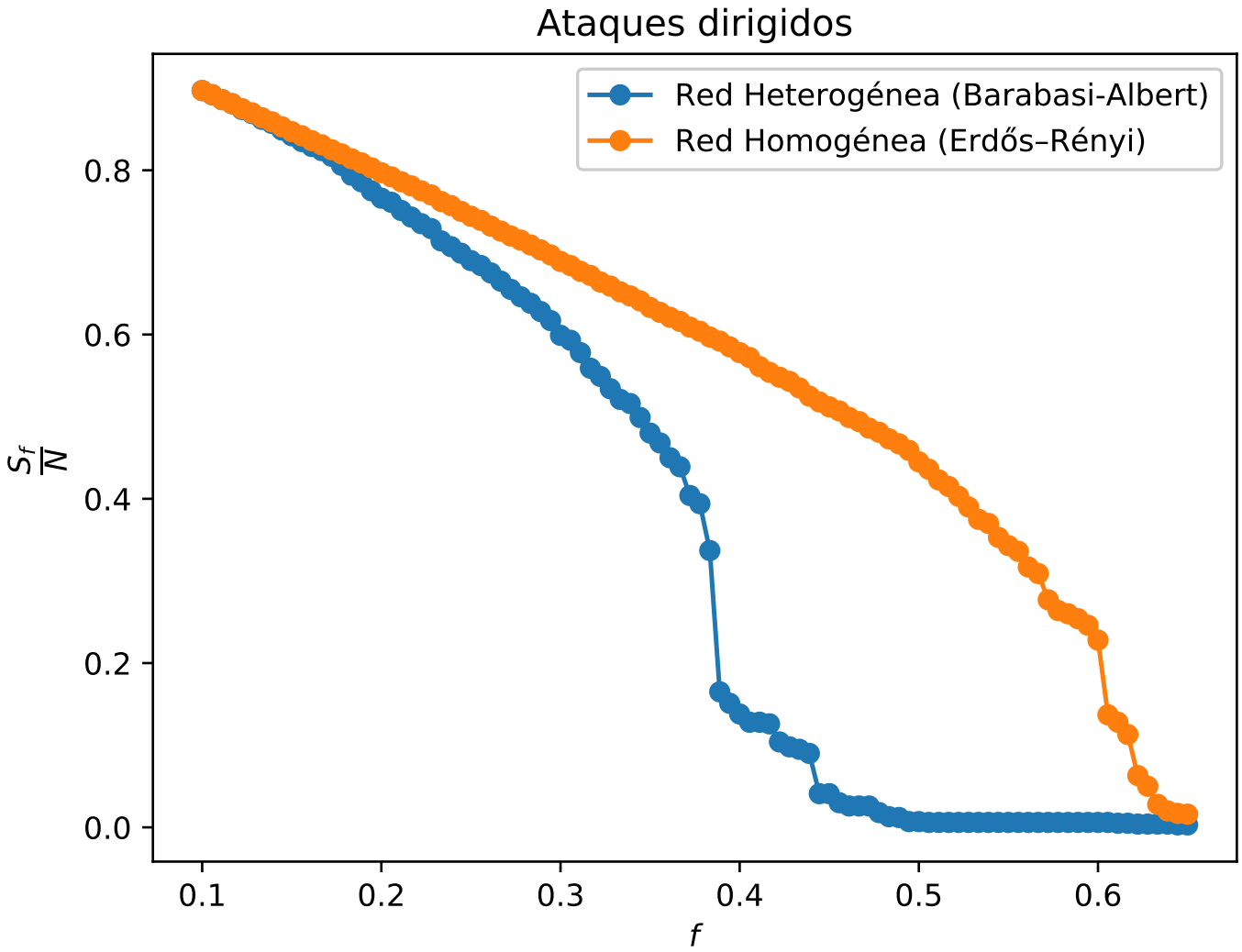


Figura: Ataques dirigidos sobre Redes Heterogéneas (Barabasi-Albert, $m = 3$) y Homogéneas (Erdős–Rényi). Número de nodos 1000

.

¿Cómo incrementamos la resiliencia de las redes?

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Designing robust scale-free networks under targeted link attack using local information

Marco Tomassini

Information Systems Institute, University of Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 29 December 2022
Received in revised form 8 February 2023
Available online 12 February 2023

Keywords:
Complex networks
Robustness of complex networks
Edge attack on scale-free networks
Vulnerability of scale-free networks
Robustness of real-world networks
Network percolation

ABSTRACT

We study the effects of deliberate attacks to important links in scale-free networks and propose simple strategies to mitigate the damage. While strategies that require global network information and heavy computation have been proposed, here we focus on heuristics that make use of local information only and are much easier to compute. Using model scale-free networks and under the constraint of an invariant degree distribution, we show by numerical simulation that our approach in the average allows to enhance the network robustness notably, as measured by the integrity of the largest connected component. Because it is fast, the approach is also applicable to real-world networks of which we give two typical examples. Overall, the results are equivalent to those obtained in more complex settings that require global network knowledge and are much more computer-intensive, an important point for large networks.

- Tomassini (2023) estudia ataques dirigidos contra las aristas de la red.
- Dichos ataques han sido menos estudiados aunque las aristas son igualmente importantes para la integridad de la red.
- Algunos ejemplos:
  - Para infligir daño a una línea eléctrica podría ser más sencillo cortar los cables entre dos torres eléctricas en lugar de destruir una planta de energía.
  - En una guerra, una ruta aérea podría ser suprimida debido a posibles ataques terrestres con misiles.
  - La destrucción de cables de comunicación en redes de comunicación.
- ¿Cómo podemos mejorar la resiliencia de la red ante ataques dirigidos a las aristas utilizando información local?

Configuración de las evaluaciones

Para responder a la pregunta, se deben tomar decisiones con respecto a:

- Modelo de red.
- Tipo de ataque.
- Medida de robustez usada para evaluar la resiliencia total de la red.

Configuración de las evaluaciones

Para responder a la pregunta, se deben tomar decisiones con respecto a:

- Modelo de red.
  - Barabási–Albert.
- Tipo de ataque.
  - Grado de la arista, $d(e)$.
  - El grado de la arista $d(e)$ de la arista $e = uv$ se define como el número vecinos de e (Yoshimoto, 2008),

$$|N(u) \cup N(v)| - 2$$

Configuración de las evaluaciones

$$e = uv$$

$$N(u) = \{a, b, c, d, e, v\}$$

$$N(v) = \{b, u\}$$

$$N(u) \cup N(v) = \{a, b, c, d, e, u, v\}$$

$$|N(u) \cup N(v)| - 2 = 5$$

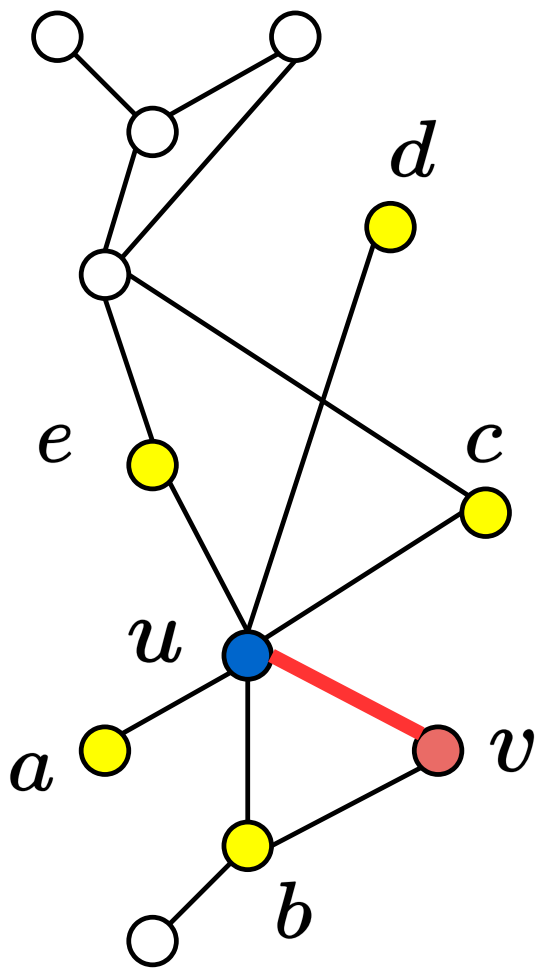


Figura: Ejemplo de cálculo de grado de la arista

Medida de robustez usada para evaluar la resiliencia total de la red

Latora and Marchiori (2001) proponen la medida de **eficiencia**:

$$\frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{l_{ij}} \tag{4}$$

donde l_{ij} es la trayectoria más corta entre los nodos i y j. Esta cantidad es finita aún en gráficas desconectadas porque se asume que para cada par de nodos desconectados $l_{ij} = \infty$.

Medida de robustez usada para evaluar la resiliencia total de la red

Otra propuesta de medida para calcular la robustez de la red durante y después de ataques es la medida R (Schneider et al., 2011).

Esta se define como el tamaño del componente conectado más grande después de una serie de ataques que remueven los nodos de forma descendente de acuerdo al grado del nodo:

$$R = \frac{1}{N} \sum_{Q=1}^{N} S(\frac{Q}{N}) \tag{5}$$

donde

- N es la cantidad de nodos en la red, Q el número de nodos removidos,
- $S(\frac{Q}{N})$ es el tamaño del componente conectado más grande después de haber removido una fracción $\frac{Q}{N}$ de nodos.

Degree-preserving edge swap for mitigating edge attacks

Reto: Dejar sin cambios el grado de cada nodo y la distribución de grados de la red completa.

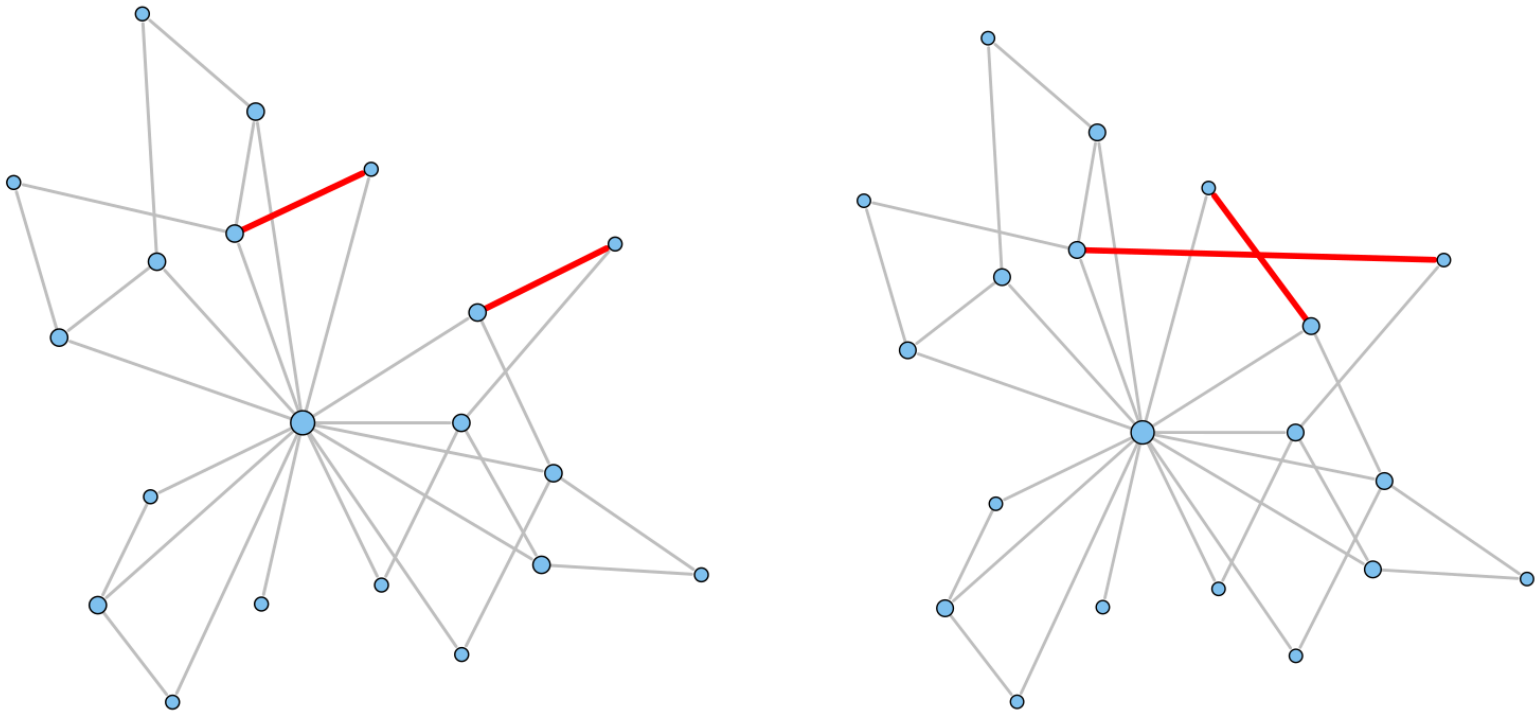


Figura: Unrestricted edge swap operation

Algorithm 1: Edge Swap Heuristic

Create graph G
Compute $R(G)$
Choose an initial temperature T
for $i = 1 \rightarrow$ number of iterations **do**
  Choose edges e_{ij}, $e_{lk} \in G$ uniformly at random
  Swap edges e_{ij}, e_{lk} obtaining a new graph H
  Compute $R(H)$
  if $R(H) > R(G)$ **then**
    $G \leftarrow H$
  else
    $G \leftarrow H$, with probability $\exp \frac{-\Delta(R)}{T}$ // $\Delta(R) = |R(H) - R(G)|$
  end if
end for

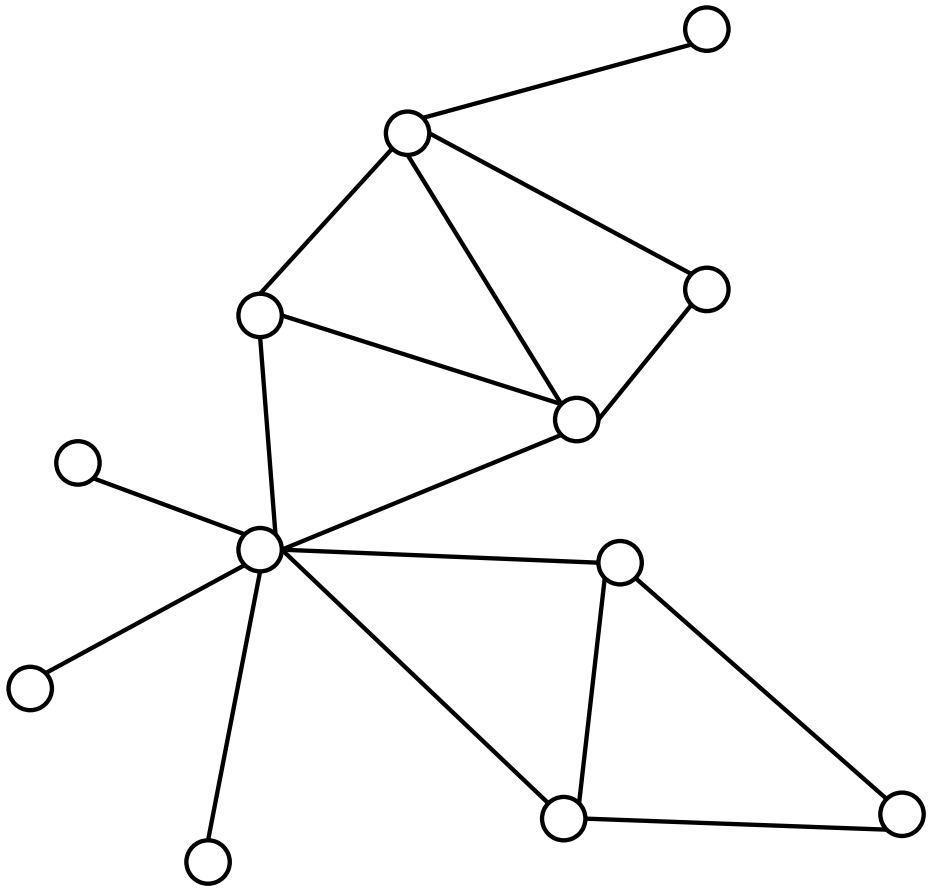
return G

- **Desventaja**:
  - El enfoque es razonable cuando se diseña una nueva red o cuando el reenrutamiento tiene un costo bajo.
  - Sin embargo, a menudo es inviable para redes reales ya existentes.
  - Ejemplos:
    - ★ Volver a cablear una línea eléctrica entre nodos distantes.
    - ★ Establecer una nueva ruta aérea a un destino arbitrario debido a limitaciones tecnológicas y económicas.
- **Alternativa**:
  - La heurística puede ser modificada de manera que las aristas son reconectadas localmente en lugar de globalmente, tal como se indica en el algoritmo 2.

Algorithm 2: Local Edge Swap Heuristic

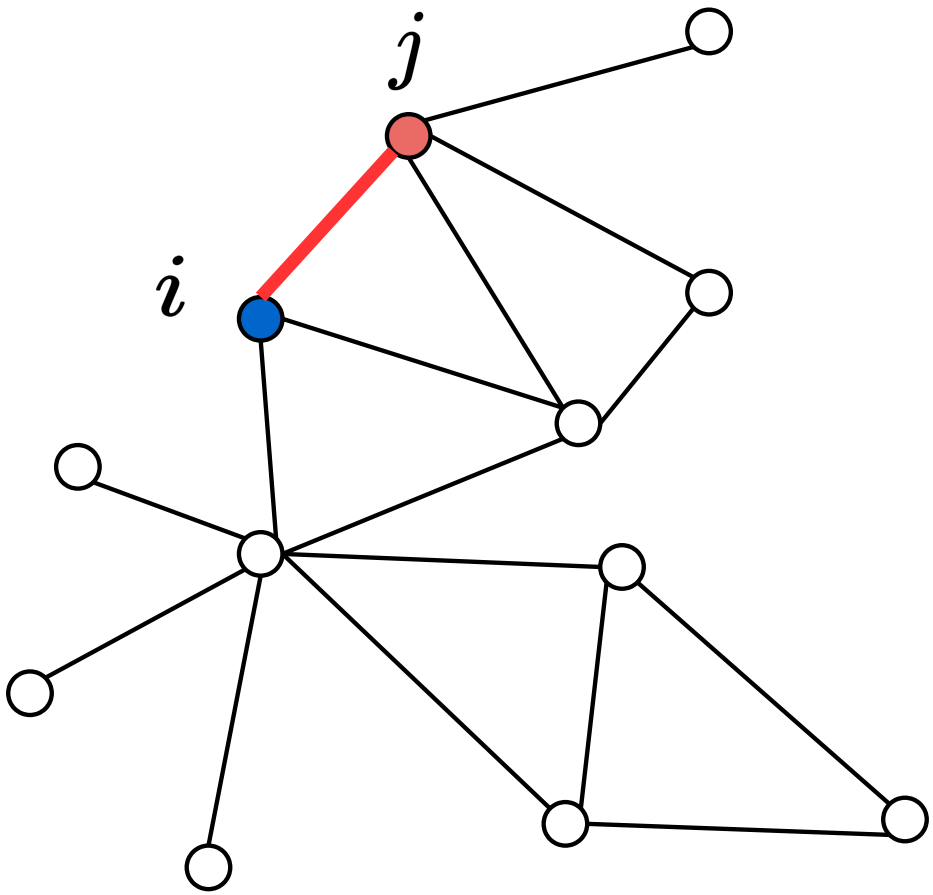
Create graph G
Compute $R(G)$
Choose an initial temperature T
for $i = 1 \rightarrow$ number of iterations **do**
  Choose edge $e_{ij} \in G$ uniformly at random
  Choose edge $e_{lk} \in G$ belonging to a neighbor of i or j
  if e_{ij} and e_{lk} are adjacent or new edges already exist **then**
    Continue with the next iteration
  else
    Swap edges e_{ij}, e_{lk} obtaining a new graph H
    Compute $R(H)$
  end if
  if $R(H) > R(G)$ **then**
    $G \leftarrow H$
  else
    $G \leftarrow H$, with probability $\exp \frac{-\Delta(R)}{T}$ // $\Delta(R) = |R(H) - R(G)|$
  end if
end for
return G

Ejemplo. Local Edge Swap Heuristic



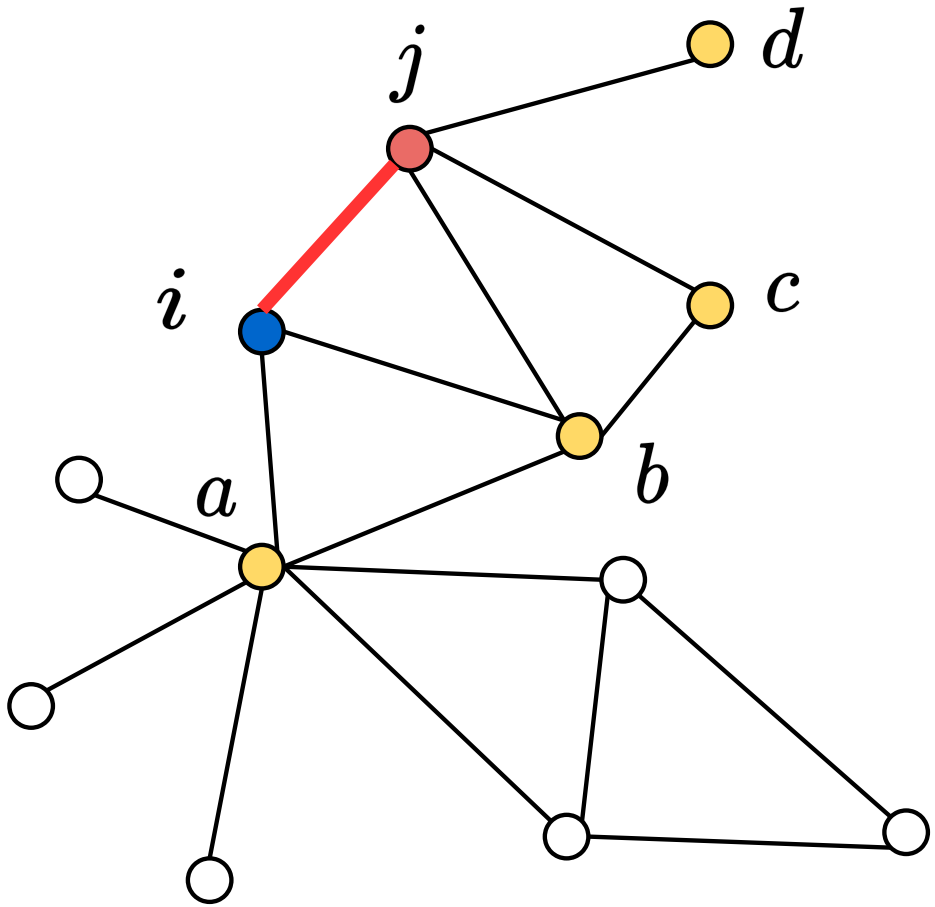
Create graph G

Ejemplo. Local Edge Swap Heuristic



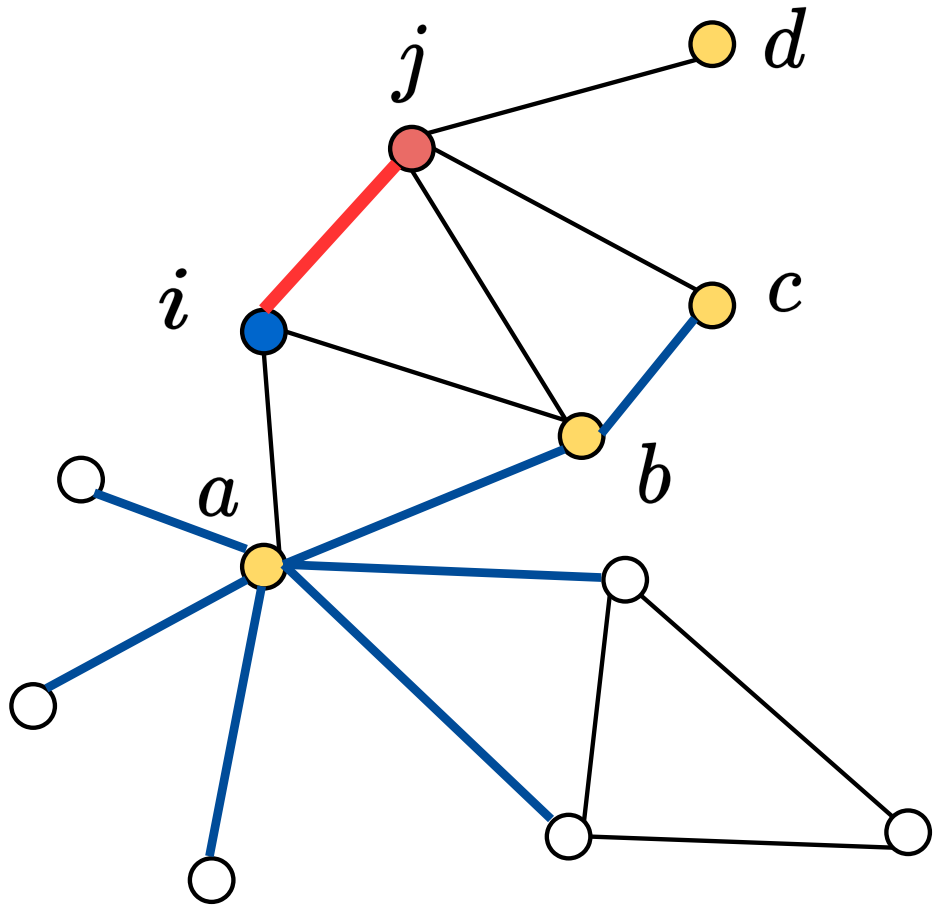
Choose edge $e_{ij} \in G$ uniformly at random

Ejemplo. Local Edge Swap Heuristic



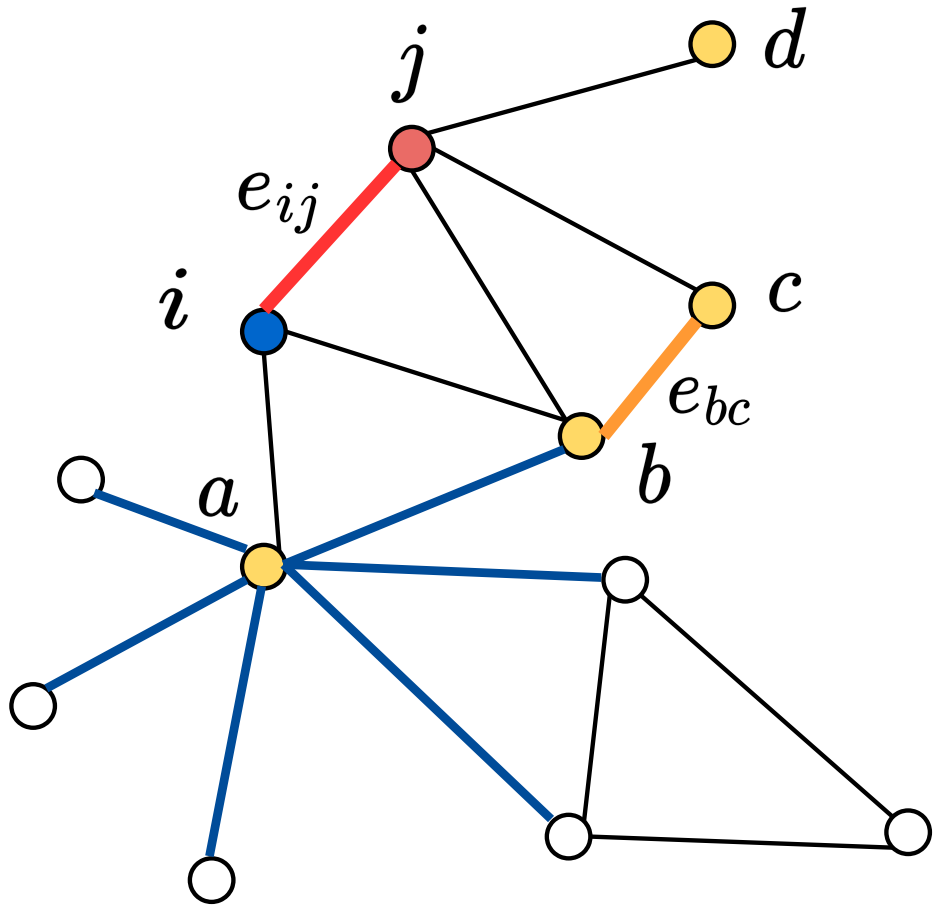
Choose edge $e_{lk} \in G$ belonging to a neighbor of i or j

Ejemplo. Local Edge Swap Heuristic



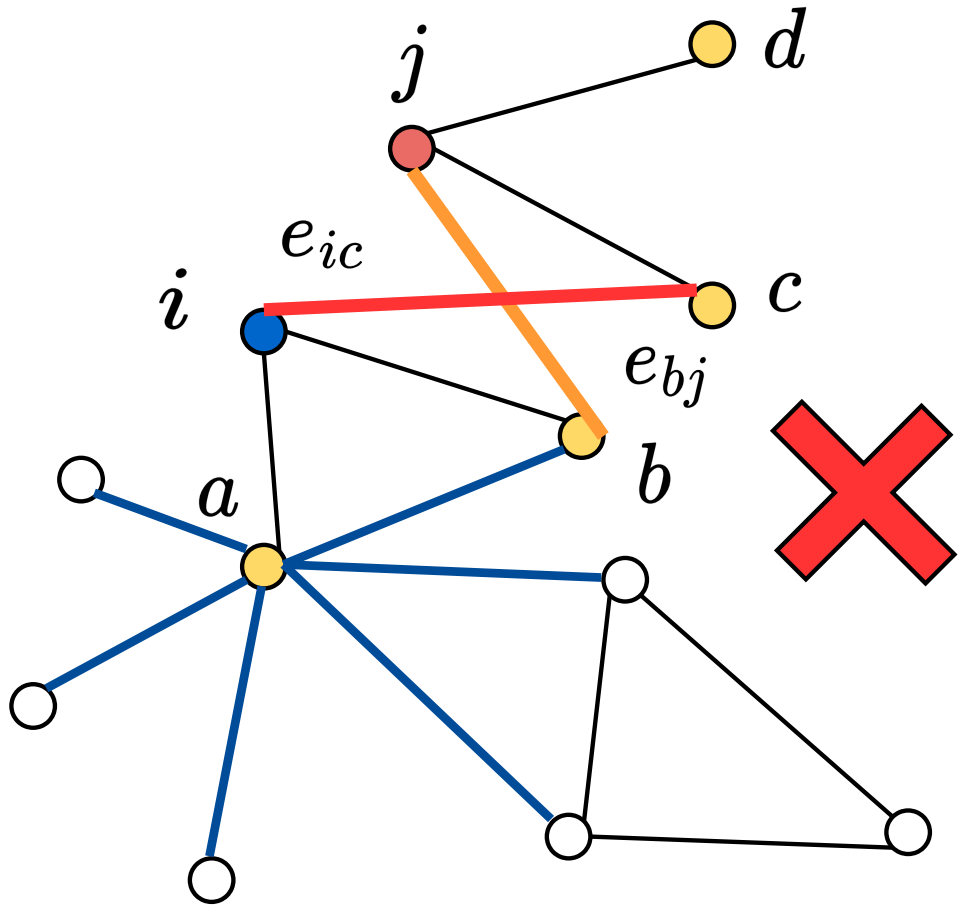
Choose edge $e_{lk} \in G$ belonging to a neighbor of i or j

Ejemplo. Local Edge Swap Heuristic



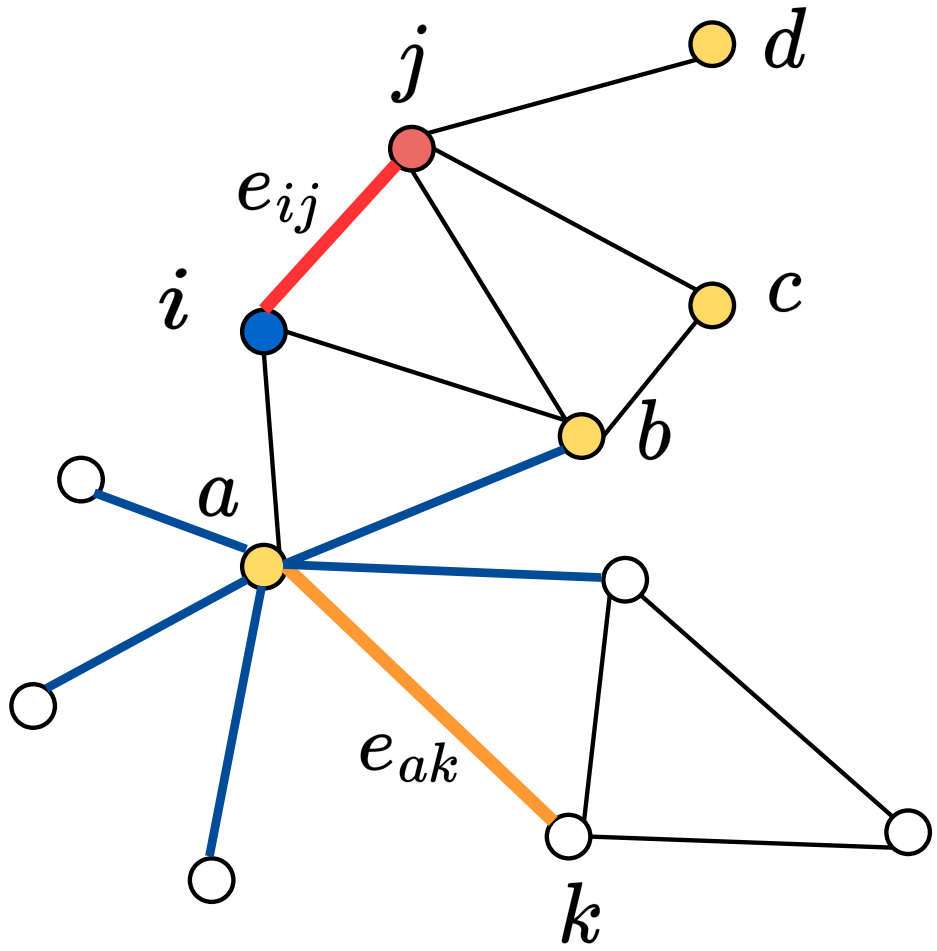
Choose edge $e_{lk} \in G$ belonging to a neighbor of i or j

Ejemplo. Local Edge Swap Heuristic



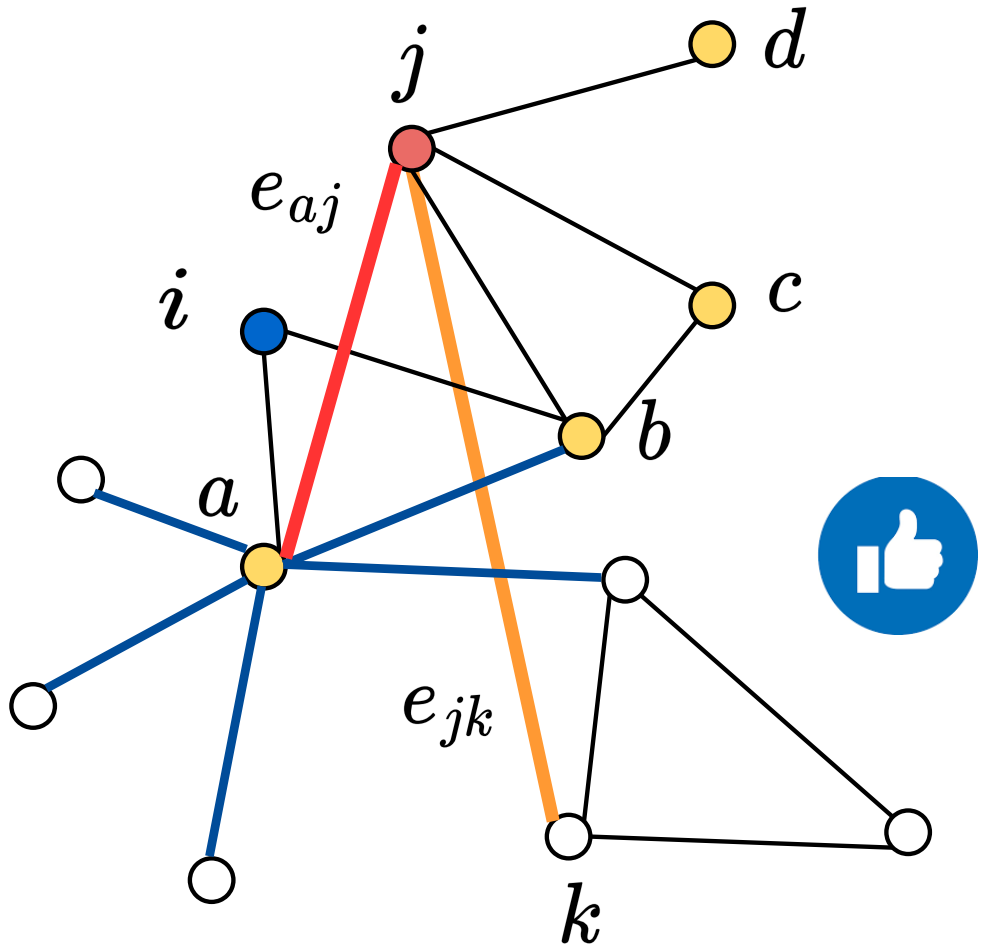
Choose edge $e_{lk} \in G$ belonging to a neighbor of i or j

Ejemplo. Local Edge Swap Heuristic



Choose edge $e_{lk} \in G$ belonging to a neighbor of i or j

Ejemplo. Local Edge Swap Heuristic



Swap edges e_{ij}, e_{lk} obtaining a new graph H

Empirical evaluation on artificial networks

- Red libre de escala del tipo Barabási-Albert con 200 nodos ($n = 200$) y 1 arista a sumar a la gráfica original ($m = 1$).
- Medida de robustez:
  - Eficiencia, $E(G)$
  - $R(G)$

Empirical evaluation on artificial networks

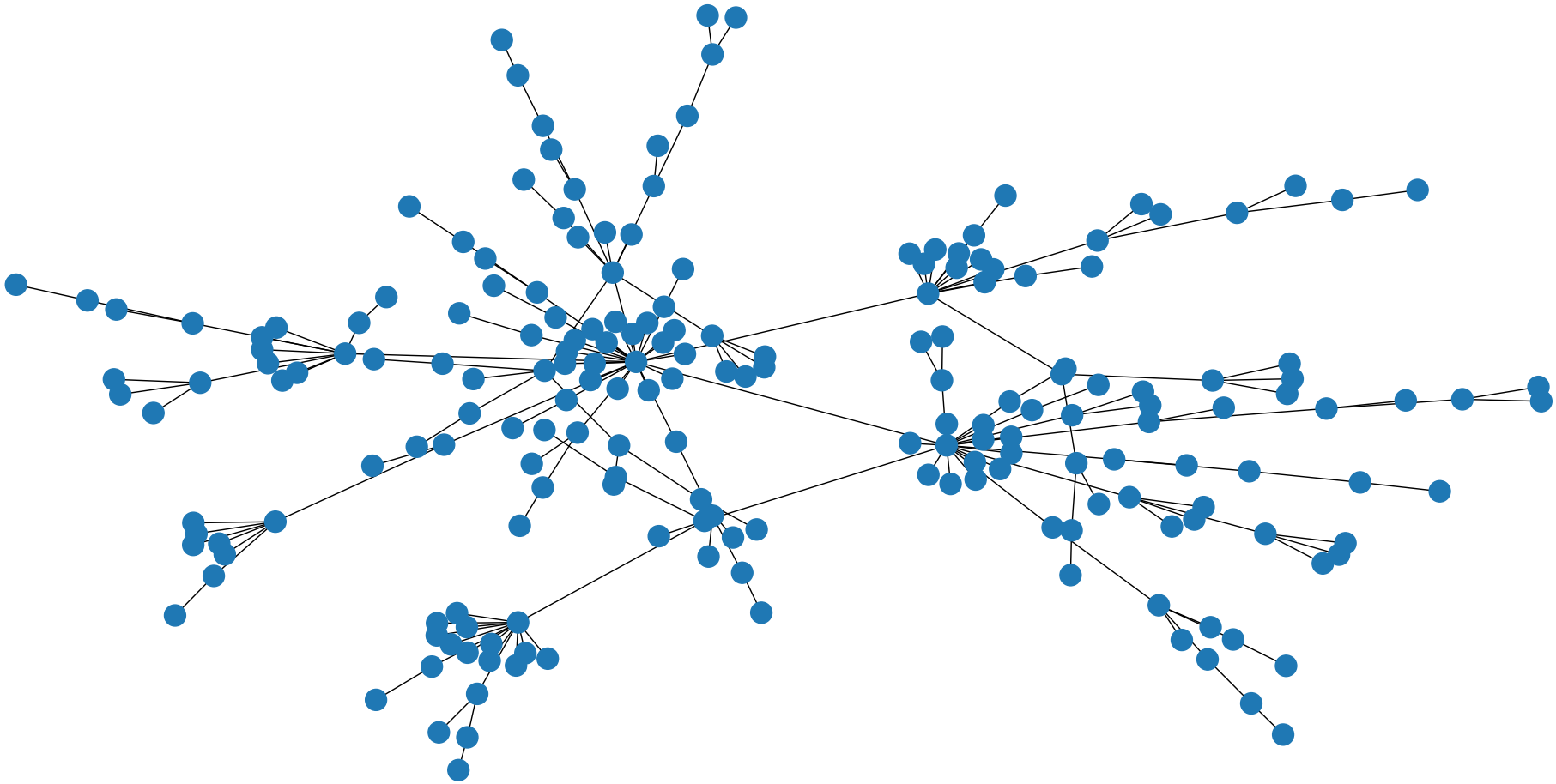


Figura: Red original

Empirical evaluation on real networks

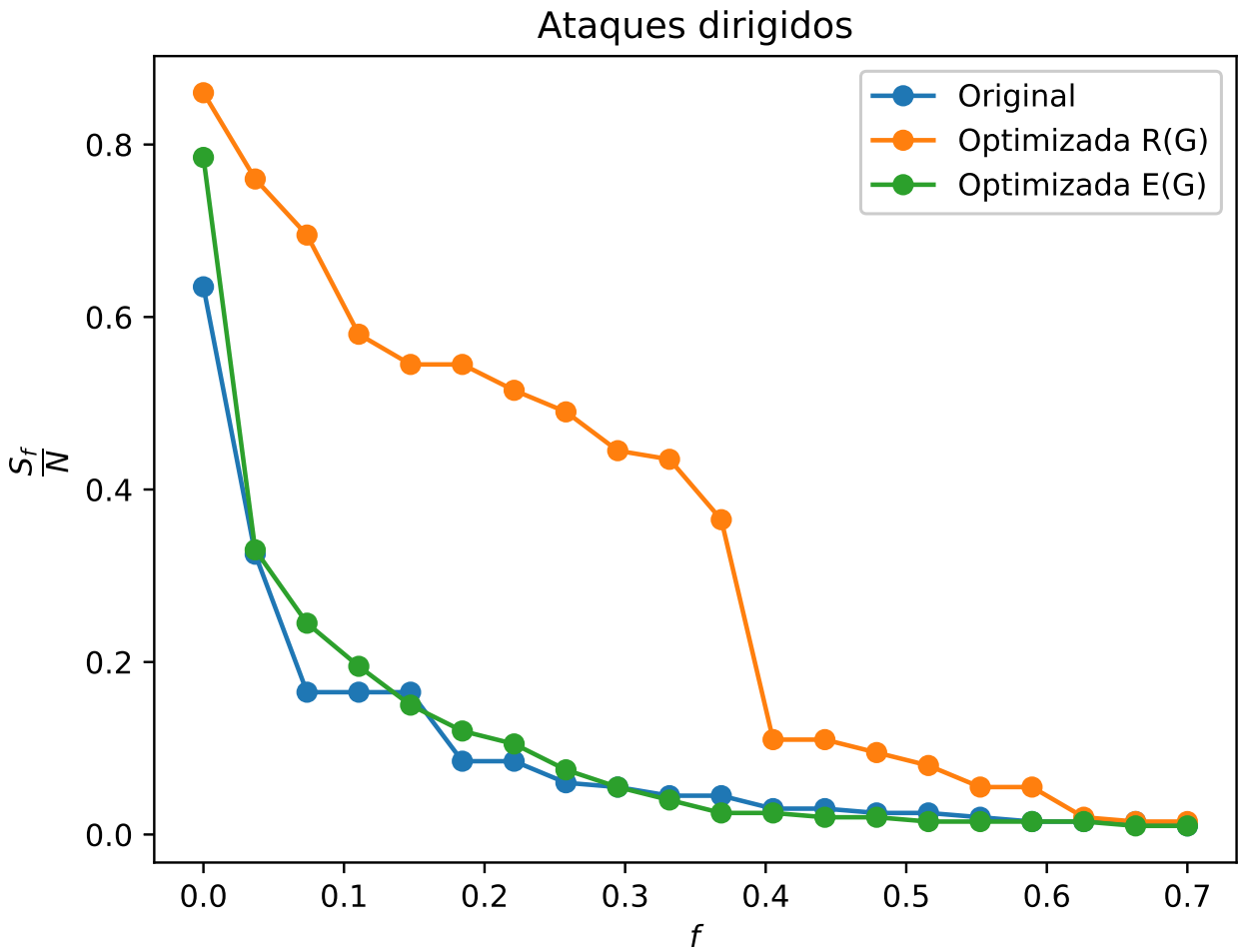


Figura: Red original vs Redes optimizadas

Empirical evaluation on artificial networks

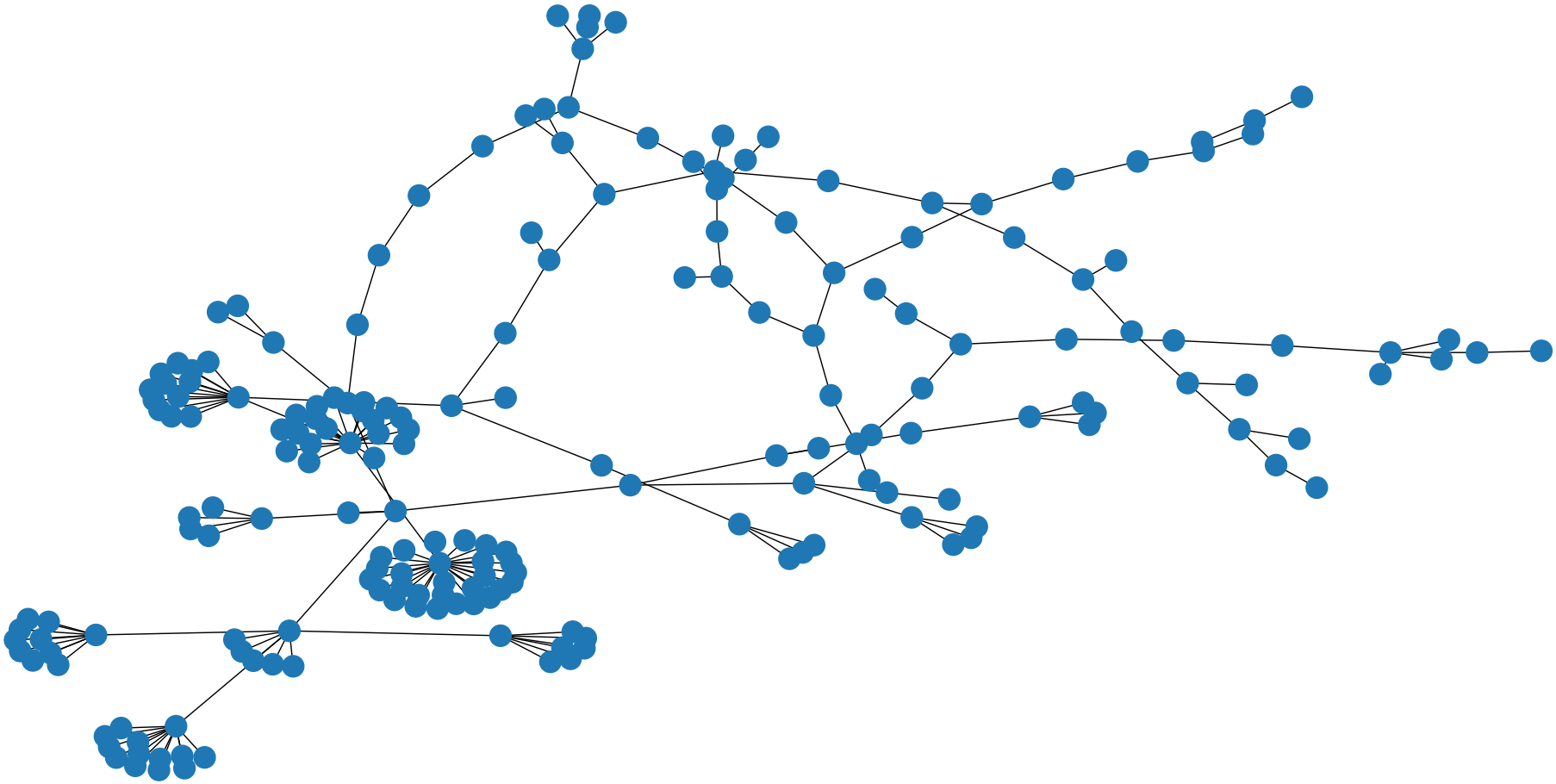


Figura: Red optimizada $R(G)$

Empirical evaluation on artificial networks

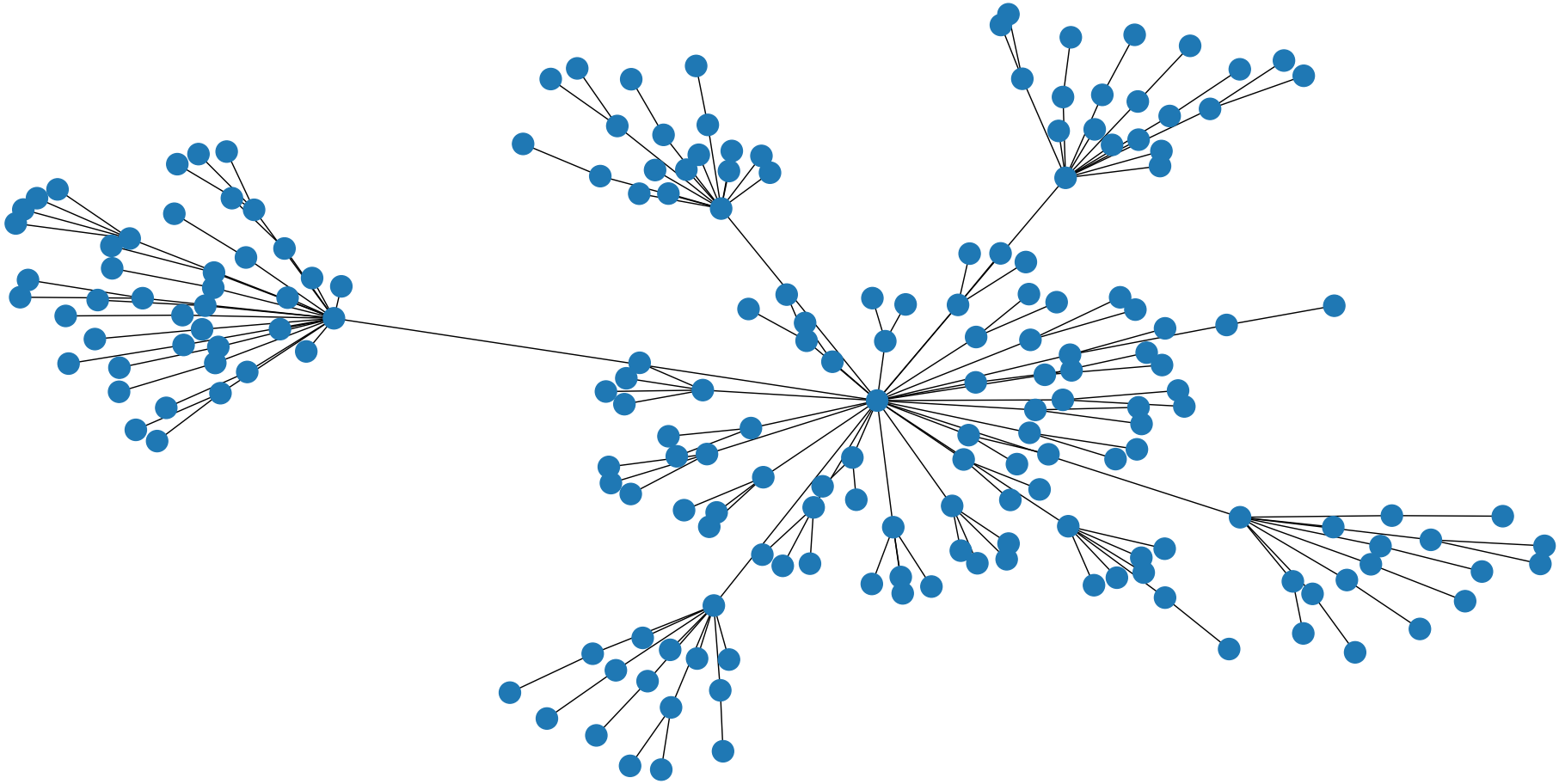


Figura: Red optimizada $E(G)$

Empirical evaluation on real networks

- We used the network representing the power grid of the western States of the U.S.A.
- The data set is from the Koblenz Network Collection KONECT (Link)[2]
- A node represents a generator, a transformator, or a substation.
- Edges represent power supply lines.
- The graph is undirected, unweighted, and has 4941 nodes and 6594 edges.
- The mean degree is 2.669.
- The degree distribution approaches a power-law with exponent $\gamma = 2{,}246$.

[2] J. Kunegis, KONECT – The Koblenz Network Collection, in: Proc. Int. Conf. on World Wide Web Companion, 2013, pp. 1343–1350.

Empirical evaluation on real networks

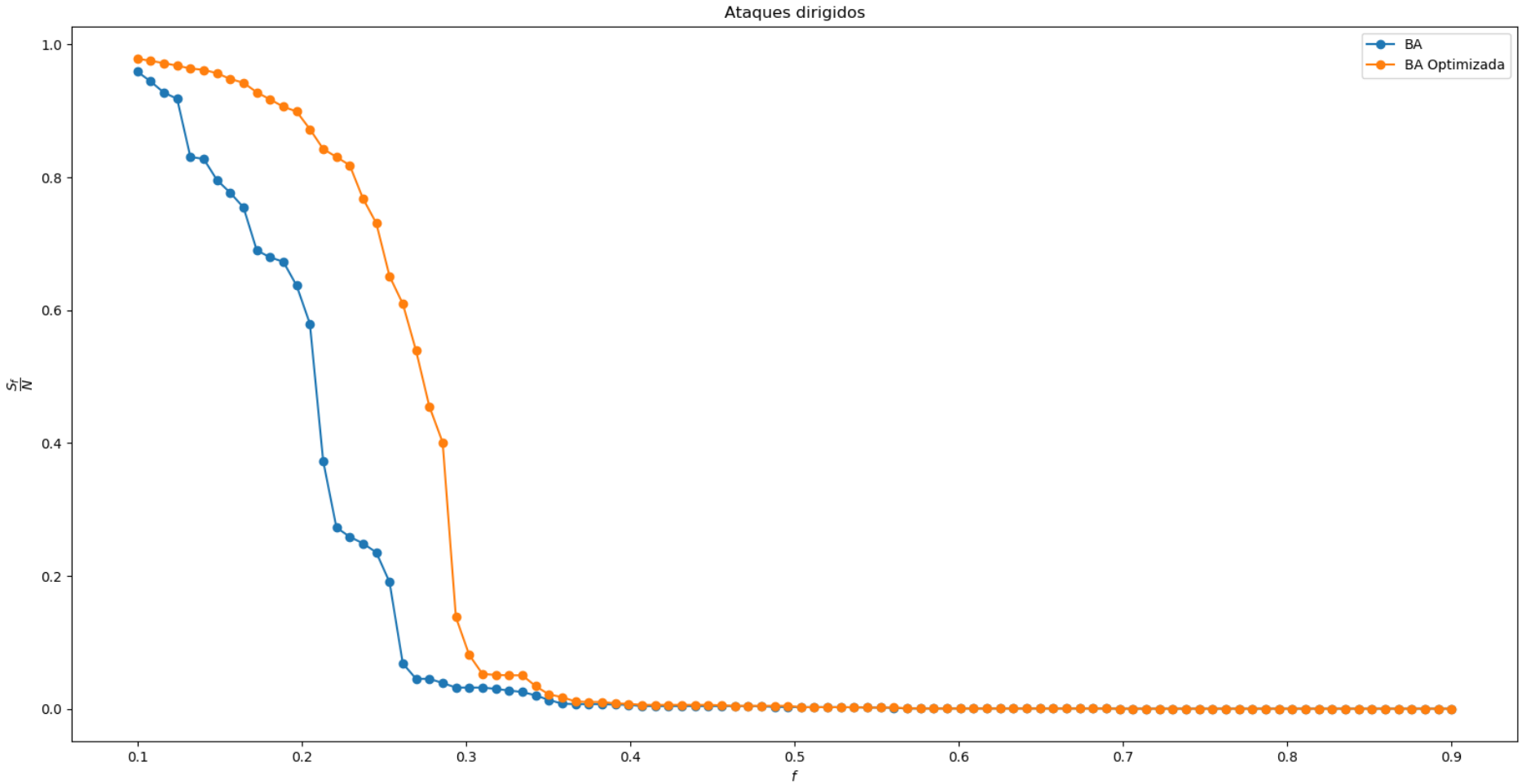


Figura: Red original vs Red optimizada

References I

Albert, R., Jeong, H., and Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *nature*, 406(6794):378–382.

Barabasi, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.

Barrat, A., Barthélemy, M., and Vespignani, A. (2008). *Dynamical Processes on Complex Networks*. Cambridge University Press.

Estrada, E. and Knight, P. A. (2015). *A first course in network theory*. Oxford University Press, USA.

Latora, V. and Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical review letters*, 87(19):198701.

Masuda, N. and Lambiotte, R. (2020). *A Guide to Temporal Networks*. WORLD SCIENTIFIC (EUROPE), 2nd edition.

Menczer, F., Fortunato, S., and Davis, C. A. (2020). *A first course in network science*. Cambridge University Press.

References II

Schneider, C. M., Moreira, A. A., Andrade Jr, J. S., Havlin, S., and Herrmann, H. J. (2011). Mitigation of malicious attacks on networks. *Proceedings of the National Academy of Sciences*, 108(10):3838–3841.

Tomassini, M. (2023). Designing robust scale-free networks under targeted link attack using local information. *Physica A: Statistical Mechanics and its Applications*, 615:128563.

Yoshimoto, K. (2008). Edge degrees and dominating cycles. *Discrete mathematics*, 308(12):2594–2599.