



ugr | Universidad
de Granada

Escuela Técnica Superior de Ingeniería
Informática y Telecomunicaciones



Redes y Sistemas Complejos
Curso 2019 - 2020

Práctica 3 :
Estudio Comparativo de Métodos
para *Poda y Visualización* de
Redes

Jesús R C

Contents

1 Primera Parte. Poda y visualización de redes con tamaño pequeño.	2
1.1 Variante de <i>Pathfinder</i> y cienciogramas escogidos	2
1.2 Análisis de los resultados obtenidos	4
1.3 Visualización de Cienciogramas con F-R y K-K	5
1.4 Comparativa de ambos Algoritmos	7
2 Segunda Parte. Comparativa tiempos de ejecución de <i>Pathfinder</i>	8
2.1 Visualización de 3 de las PFNETs obtenidas	9

1 Primera Parte. Poda y visualización de redes con tamaño pequeño.

1.1 Variante de *Pathfinder* y cienciogramas escogidos

Para la realización de esta primera parte he elegido ha sido *Binary Pathfinder*, por su mayor eficiencia.

Y las redes seleccionadas (sin ningún criterio especial) han sido :

1. Spain (2002)
2. USA (2002)
3. China (2002)
4. Japan (2002)
5. Germany (2002)
6. World (2002)
7. Spain (2004)
8. Cuba (2004)
9. Argentina (2005)
10. Canada (2005)

Spain-2002		$r = \infty$
$(n = 264)$	Nº enlaces/Densidad	Dist.Media
Original	21807/0.628	1.372
2	320/0.009	4.7
3	274/0.008	5.828
4	265/0.008	6.546
5	263/0.008	6.870
263	263/0.008	6.870

USA-2002		$r = \infty$
$(n = 276)$	Nºenlaces/Densidad	Dist.Media
Original	31292/0.825	1.175
2	314/0.008	5.226
3	287/0.008	6.294
4	279/0.007	7.531
5	277/0.007	7.603
275	275/0.007	8.168

China-2002		$r = \infty$
(n = 259)	Nºenlaces/Densidad	Dist.Media
Original	20661/0.618	1.383
2	306/0.009	4.983
3	260/0.008	7.593
4	268/0.008	5.986
5	262/0.008	6.823
258	260/0.008	7.593

Japan-2002		$r = \infty$
(n = 265)	Nºenlaces/Densidad	Dist.Media
Original	21754/0.622	1.379
2	316/0.009	4.972
3	267/0.008	7.134
4	279/0.008	5.689
5	269/0.008	6.271
264	267/0.008	7.134

Germany-2002		$r = \infty$
(n = 269)	Nºenlaces/Densidad	Dist.Media
Original	25395/0.705	1.295
2	313/0.009	5.181
3	269/0.007	7.908
4	277/0.008	6.008
5	272/0.008	6.847
268	270/0.007	7.698

World		$r = \infty$
(n = 218)	Nºenlaces/Densidad	Dist.Media
Original	20154/0.852	1.147
2	280/0.012	4.938
3	217/0.009	7.697
4	233/0.010	6.235
5	223/0.009	6.937
217	220/0.009	7.190

Spain-2004		$r = \infty$
(n = 269)	Nºenlaces/Densidad	Dist.Media
Original	24991/0.693	1.307
2	332/0.009	4.775
3	270/0.007	7.712
4	280/0.008	5.880
5	272/0.008	7.034
268	271/0.008	7.198

Cuba-2004		
		$r = \infty$
(n = 237)	Nºenlaces/Densidad	Dist.Media
Original	9518/0.340	1.694
2	285/0.010	4.406
3	247/0.009	5.975
4	255/0.009	5.271
5	249/0.009	5.606
236	247/0.009	5.975

Argentina-2005		
		$r = \infty$
(n = 266)	Nºenlaces/Densidad	Dist.Media
Original	17938/0.509	1.497
2	324/0.009	4.700
3	267/0.008	6.211
4	277/0.008	5.726
5	269/0.008	6.079
265	268/0.008	6.113

Canada-2005		
		$r = \infty$
(n = 288)	Nºenlaces/Densidad	Dist.Media
Original	30574/0.740	1.260
2	340/0.0087	4.993
3	287/0.007	7.789
4	301/0.007	5.931
5	290/0.007	6.734
287	287/0.007	7.789

1.2 Análisis de los resultados obtenidos

Por norma general en todos los cienciogramas es una disminución considerable del número de enlaces entre la *Red Original* y una primera poda con $q = 2$, que, como se puede ver, es bastante fuerte, ya que elimina la gran mayoría de enlaces, por ejemplo, en el cienciograma de los España en el año 2002 pasamos de 21807 enlaces a 320, es decir, tras la poda sólo quedan el 1.467%.

Una vez hecha la primera poda, las siguientes son bastante más suaves, por tanto, podemos concluir que **cuanto más bajo sea el valor de q , mayor va a ser la Poda**.

Llega un momento en el que ya no se puede podar más. En la mayoría de cienciogramas estudiados, este momento llega cuando $q = 5$, aunque no siempre es así, porque hay ejemplos en los que desde $q = 5$ hasta $q = n - 1$ podan varios enlaces más:

- World
- United States 2002
- Argentina 2005, etc.

Por tanto, podemos concluir que **A partir de, más o menos, $q = 5$ la PFNET($r = \infty, q$) alcanza su valor muy próximo a su valor máximo de poda**, además, **no hay mucha dispersión en ese valor**, ya que casi todos los cienciogramas llegan al límite ahí.

Otro detalle a tener en cuenta es que conforme se va podando la red, va reduciéndose la *Densidad* de la misma, ya que esta depende del número de enlaces, L , de la red, y que al ir eliminando enlaces al podar, también la reducimos.

1.3 Visualización de Cienciogramas con F-R y K-K

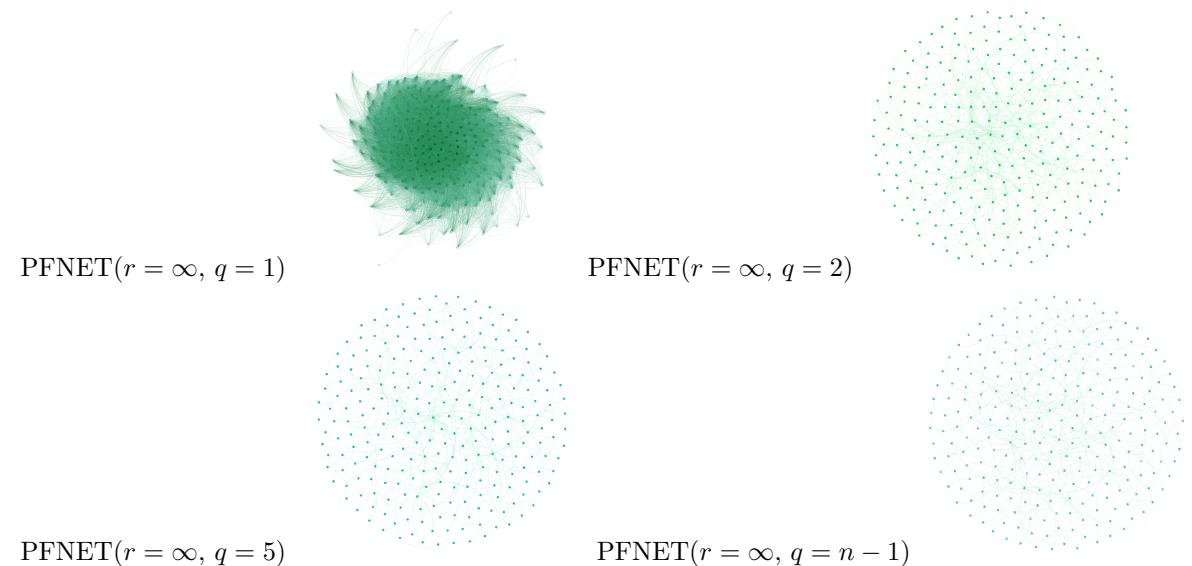


Figure 1: Visualizaciones de las PFNETs del cienciograma Spain-2002 realizadas con el método de distribución **Fruchterman-Reingold**

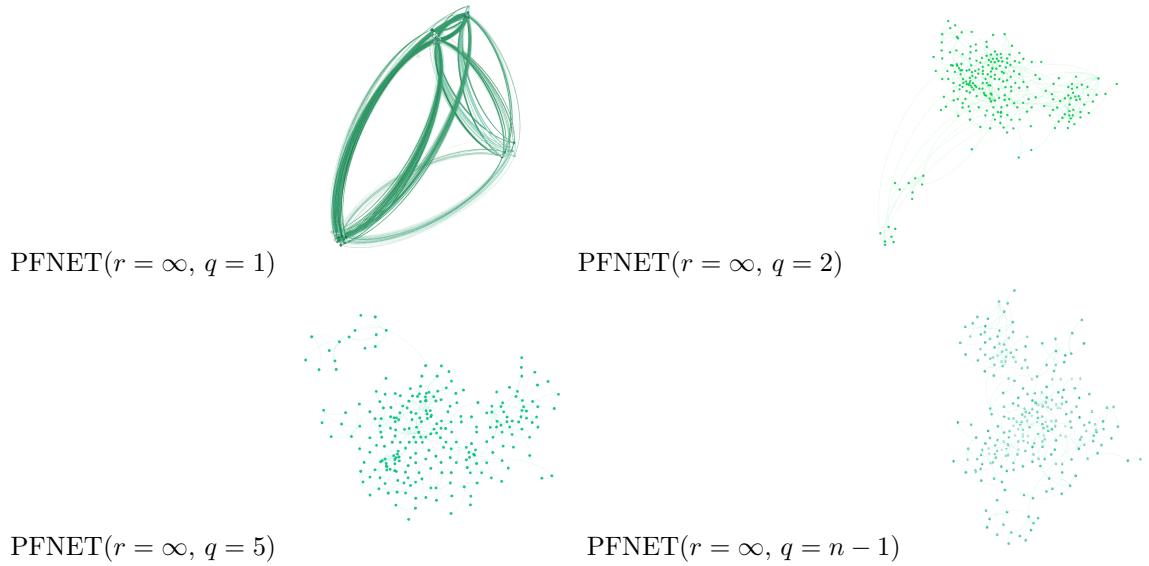


Figure 2: Visualizaciones de las PFNETs del cienciograma Spain-2002 realizadas con el método de distribución **Kamada & Kawai**

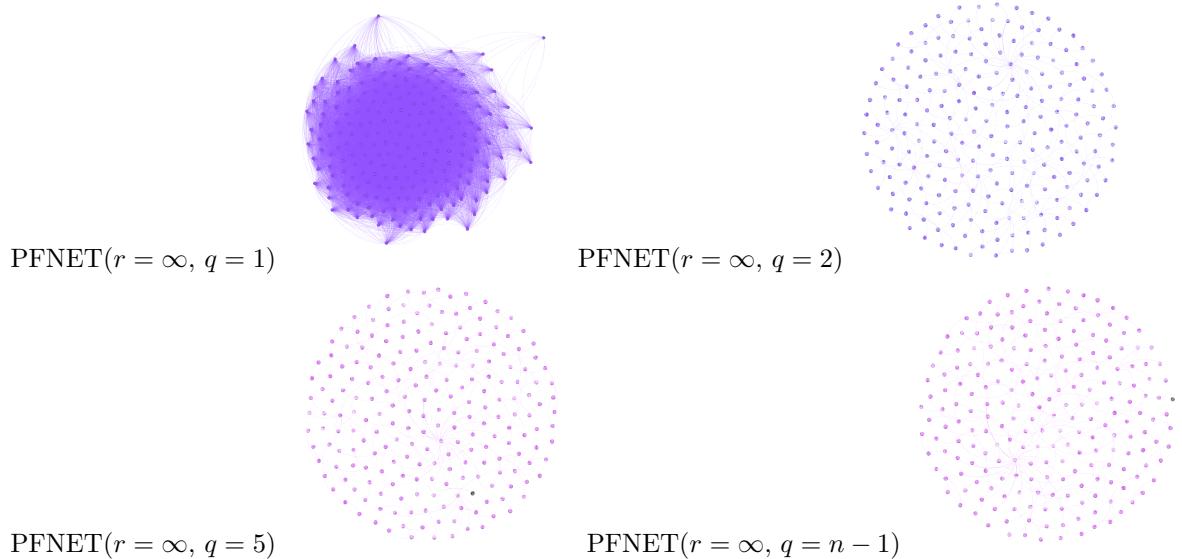


Figure 3: Visualizaciones de las PFNETs del cienciograma World-2002 realizadas con el método de distribución **Fruchterman-Reingold**

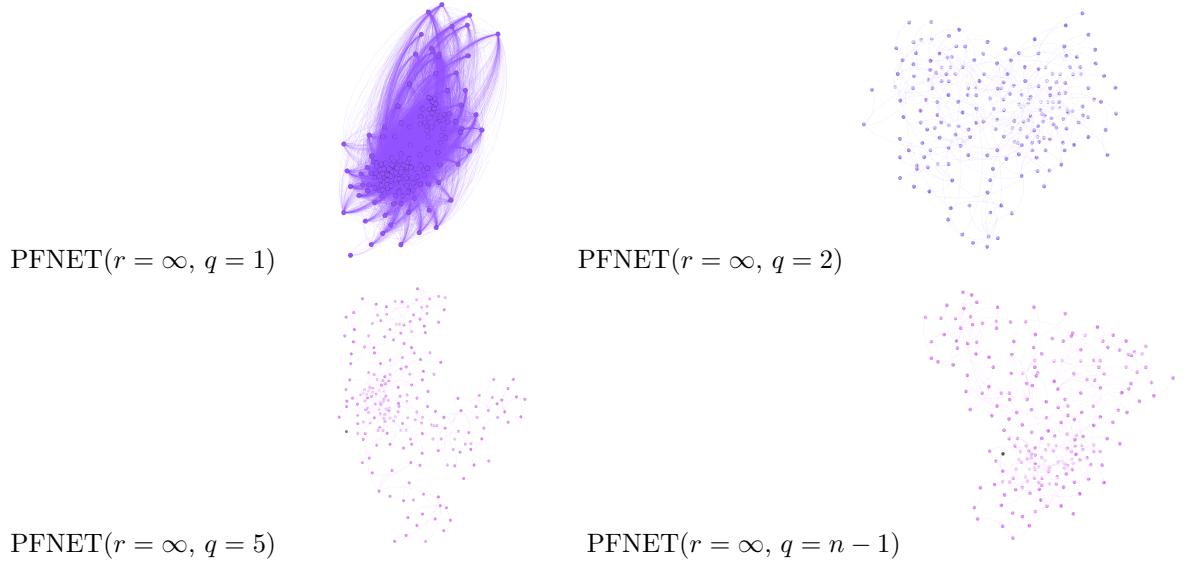


Figure 4: Visualizaciones de las PFNETs del cienciograma World-2002 realizadas con el método de distribución **Kamada & Kawai**

1.4 Comparativa de ambos Algoritmos

Para **Fruchterman-Reingold** en ambas redes obtenemos un resultado muy parecido: una red circular donde los nodos están a la misma distancia los unos de los otros.

Otro aspecto a destacar de ambas figuras es la gran diferencia en cuanto a poda desde $q = 2$ hasta $q = n - 1$: la visualización se queda mucho más “limpia” en la última poda. Con esta visualización se nota bastante la diferencia de enlaces. Por último, esta visualización impide ver las comunidades científicas que hay y además tampoco permite ver cómo de “alejadas” están dos disciplinas científicas.

Para **Kamada & Kawai** con $q = 2$ los nodos centrales forman una comunidad en el centro mientras que con $q = n - 1$ (y conforme va aumentando la poda) se van formando más comunidades y se acentúa la representación.

También se nota mucho la diferencia de poda entre la red con $q = 2$ y la red con $q = n - 1$, en este caso incluso mucho más acentuada debido a que se diferencian distintas comunidades cuando eliminamos enlaces.

Entre ambas visualizaciones, la que más me ha gustado ha sido la de **Kamada & Kawai**, ya a que es la visualización que nos permiten distinguir las disciplinas científicas más importantes, debido a que están colocadas en la parte central de la red, y además también nos permite ver cómo de alejadas están dos disciplinas científicas.

2 Segunda Parte. Comparativa tiempos de ejecución de *Pathfinder*

n	Media $ E $	Media D	Original	Binary	Fast	MST (baja compl.)	MST (práctico)
500	1252	0.01	82.937	6.3310	0.4045	0.0374	0.0065
1000	4999	0.01	1304.4016	89.3710	2.9496	0.0201	0.0203
2000	20041	0.01	> 1800	1234.494	23.625	0.156	0.0792
5000	124525	0.00998	> 1800	> 1800	361.3300	0.4194	0.455
10000	498670	0.00998	> 1800	> 1800	> 1800	7.260	5.7310

Table 1: Comparación de los tiempos de ejecución de las variantes de Pathfinder en redes aleatorias

Las redes no son demasiado densas. En general, la densidad ha sido alrededor de $D = 0.01$ de media y el número de enlaces medio está muy por debajo de L_{max} debido a la dispersión de las redes generadas.

Se aprecia la diferencia de tiempo entre el algoritmo *MST* y el resto. Esta diferencia se acentúa conforme se aumenta el tamaño de la red, en el caso de mayor n el *Tiempo PF original* tarda entre 30 y 40 minutos, mientras que el *MST*, sólo un segundo.

Aunque, en teoría, exista poca diferencia de eficiencia entre el *MST* y el *Binary*, en cuanto crece el tamaño de la red de 5000 nodos en adelante, el *Binary* queda en evidencia.

El algoritmo *Fast*, consigue una mejora que se hace notar en los tiempos de ejecución, el *Binary* es mucho más lento que el *Fast*. Además, en el *Fast* el tiempo de ejecución crece mucho más lento conforme aumenta la complejidad llegando a aguantar hasta redes con 1000 nodos en un tiempo considerable.

Por último, las dos versiones diferentes de *MST*: A pesar de que el algoritmo teórico tiene una mejor eficiencia que el práctico ($O(n^2 \log(n))$ frente a $O(n^3)$) esta diferencia hace que el algoritmo práctico sea más rápido, y se nota cada vez más conforme crece el tamaño de la red, llegando a obtener una diferencia de dos segundos entre ambos algoritmos para 10000 nodos.

2.1 Visualización de 3 de las PFNETs obtenidas

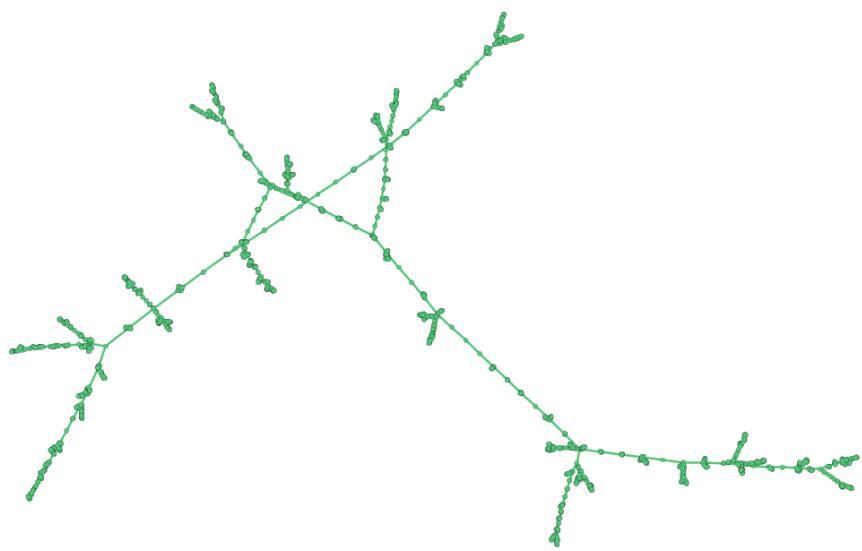


Figure 5: Visualización de la PFNET($r = \infty$, $q = n - 1$) de la Red Aleatoria realizada con el método de distribución *Fast-Pathfinder*. La PFNET tiene **n=2000** y $|E| = < 1999 >$

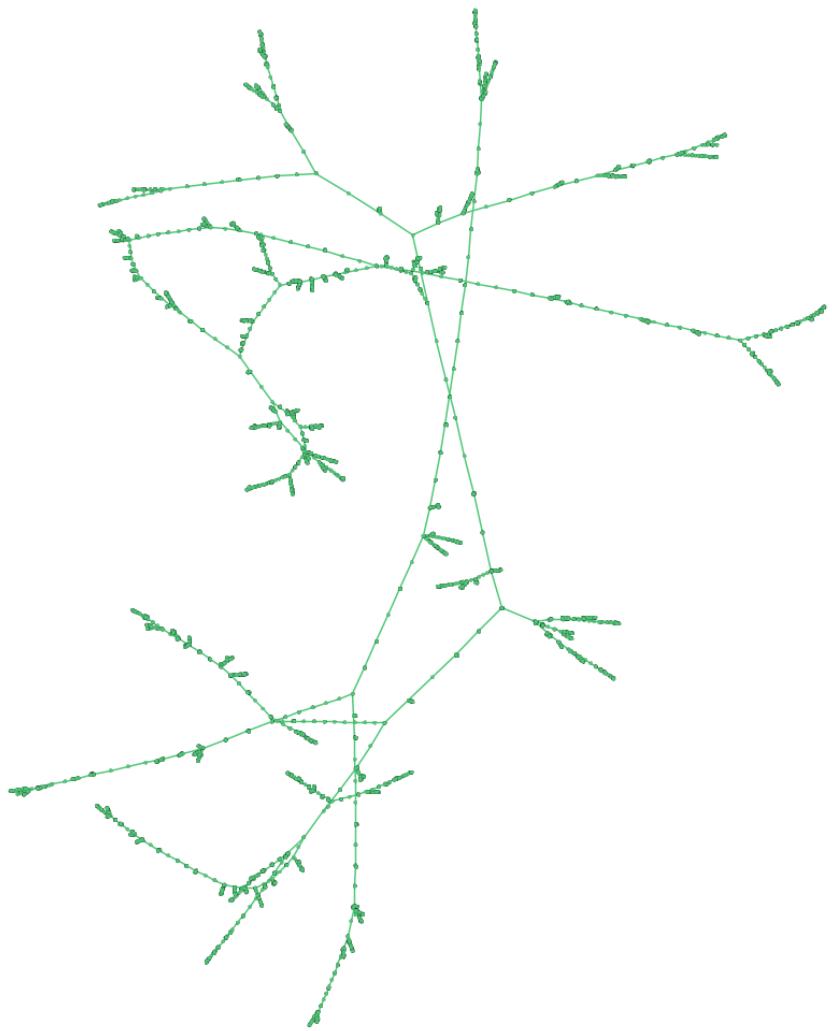


Figure 6: Visualización de la PFNET($r = \infty$, $q = n - 1$) de la Red Aleatoria realizada con el método de distribución *MST-PF-Theoretical*. La PFNET tiene **n=5000** y $|E| = < 5000 >$

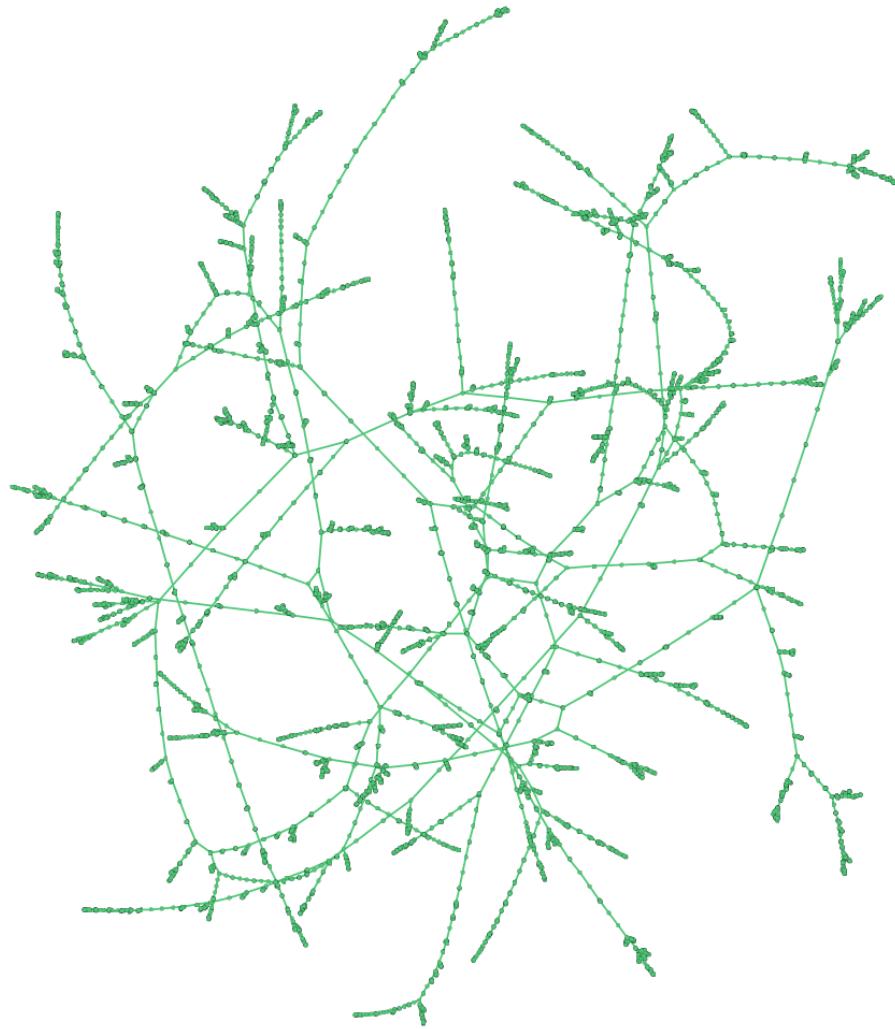


Figure 7: Visualización de la PFNET($r = \infty$, $q = n - 1$) de la Red Aleatoria realizada con el método de distribución *MST-PF-Practical*. La PFNET tiene **n=10000** y $|E| = < 10006 >$