

Grafos de Vecindad Relativa en la esfera

Estudio y Modelado

Miguel Ángel Zamorano Presa

Tecnologías para la Información en las Ciencias,UNAM, ENES Morelia

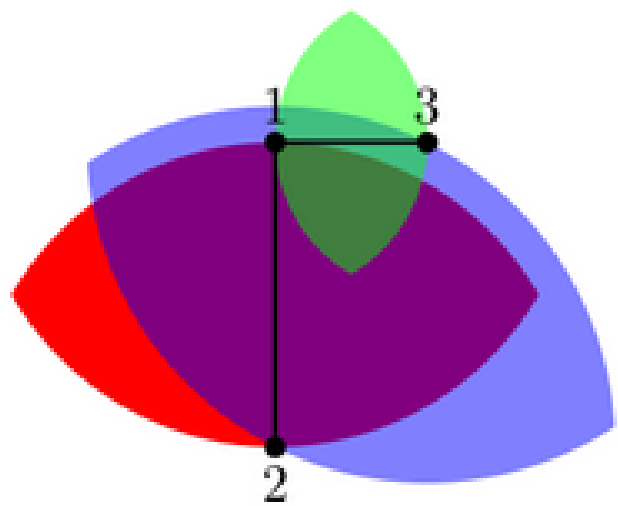
mikezpresa@comunidad.unam.mx



¿Qué son las RNG?

Los grafos de vecindad relativa (RNG, Relative Neighborhood Graphs) fueron introducido por Toussaint en 1980 [4].Se definen en término de una relación geométrica entre los puntos: una arista es compartida por dos puntos si y solo si se cumple que la distancia entre u y v es menor o igual a la distancia máxima entre cada uno de estos puntos y cualquier otro punto z del grafo:

$$d_{uv} \leq \max\{d_{uz}, d_{vz}\}$$



Las RNG han encontrado aplicación en diversos campos: Modelado 3D, renderización,en técnicas para minería de datos ,por mencionar algunos.

Limitaciones de las RNG

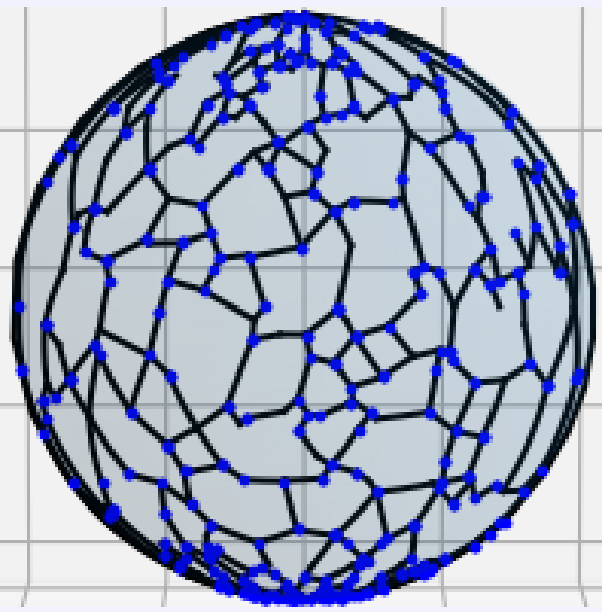
- Identificar una RNG es computacionalmente demandante
- En la literatura solo se han estudiado RNG en planos sin curvaturas

Hipótesis de trabajo

- Es posible reducir la complejidad de las RNG de manera aproximada

$$\text{no RNG} \simeq \text{RNG}$$

- Es posible extender las RNG a superficies con curvatura



Objetivos

Objetivo General

Estudiar propiedades estadísticas de RNGs sobre la superficie de la esfera unitaria y proponer grafos random que aproximen las RNGs.

Objetivos específicos

1. Extraer las propiedades estadísticas relevantes de las RNG en la esfera unitaria (distribución de las aristas, grado promedio, correlaciones, etc.)
2. Proponer un modelo que genere grafos random que aproxime una RNG utilizando nuestros resultados del análisis estadístico en forma de heurísticas (distribución de número de links y distribución del valor de distancia de links).

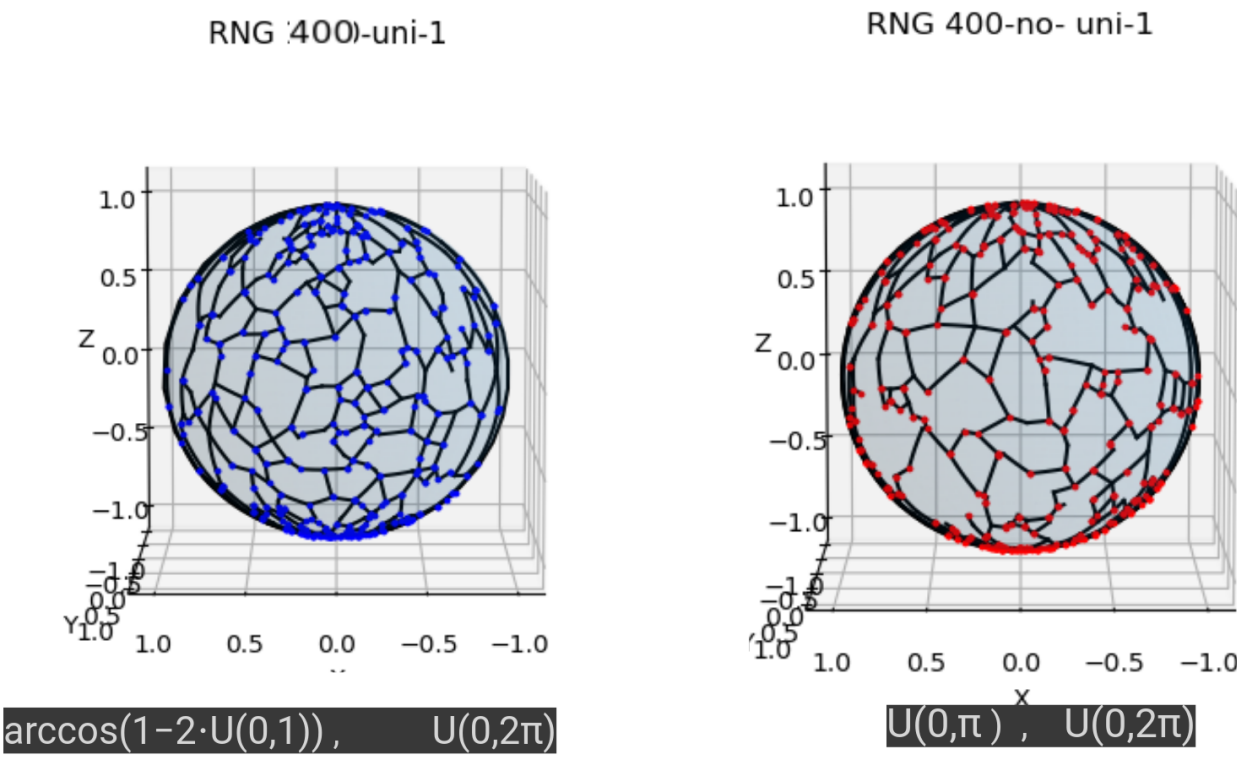
Metodología

Fase 1

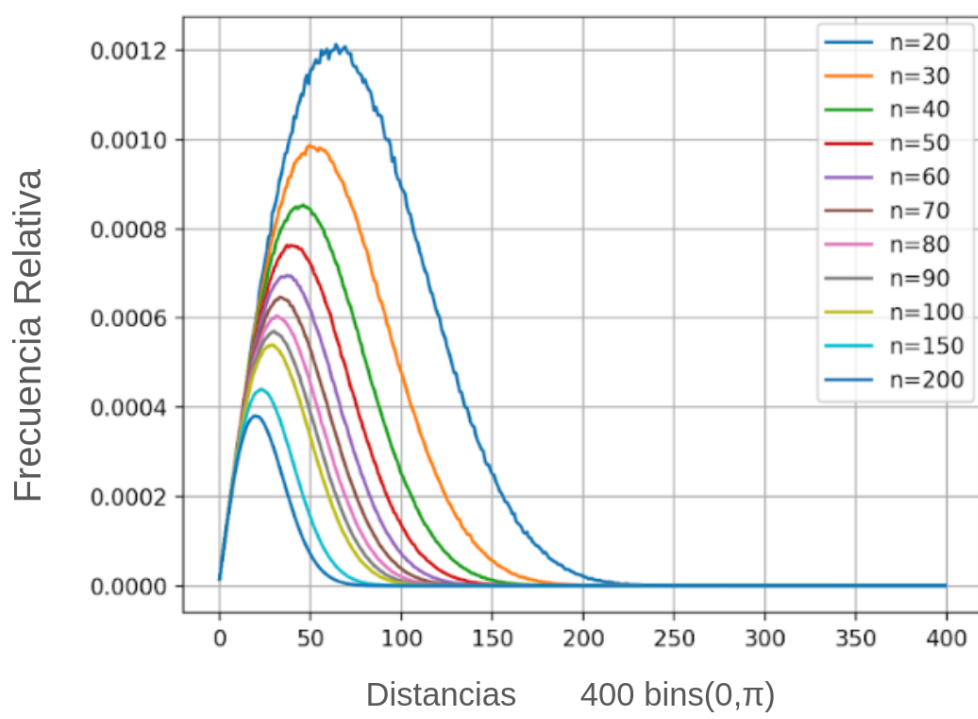
Modelado y Simulación Inicial

1. Implementamos y aplicamos el algoritmo original propuesto en [4] para la construcción de las RNG compatible con la superficie esférica.

2. construimos su respectivo recolector de resultados y un visualizador



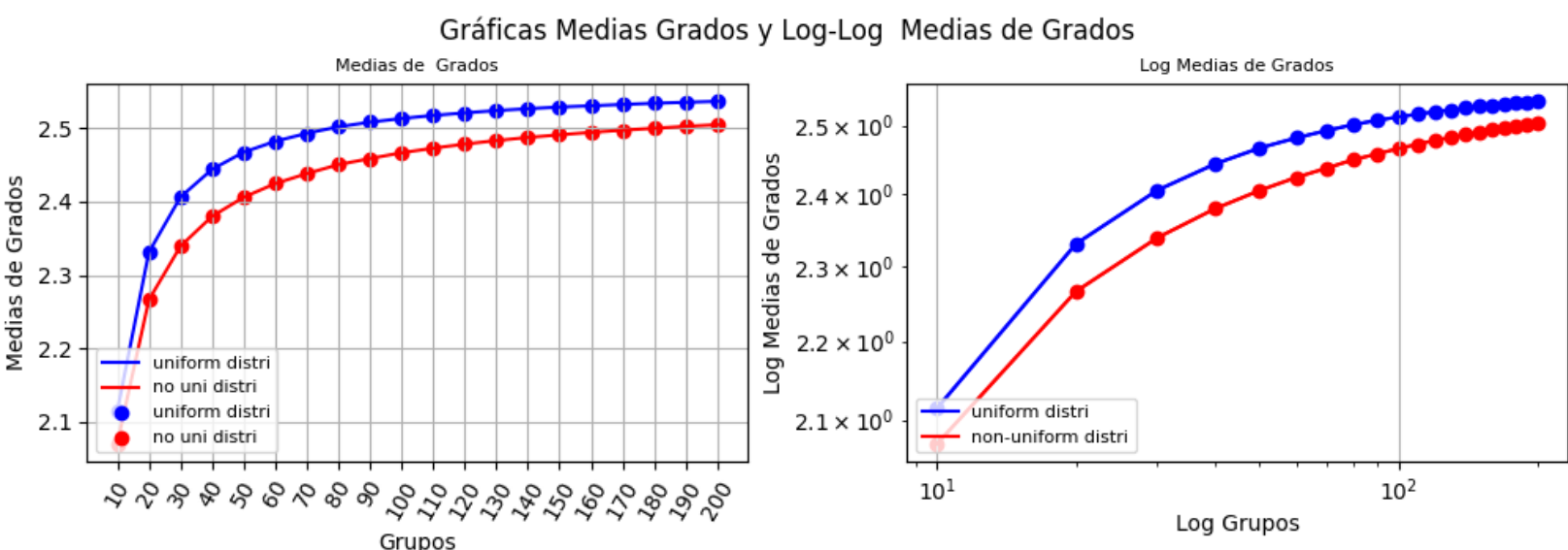
3. Organizamos nuestros grafos generados en 2 categorías,de acuerdo al tipo de distribución utilizada para generar los puntos aleatorios.Por cada categoría, generamos 20 grupos,cada grupo con un total de 10^5 grafos.



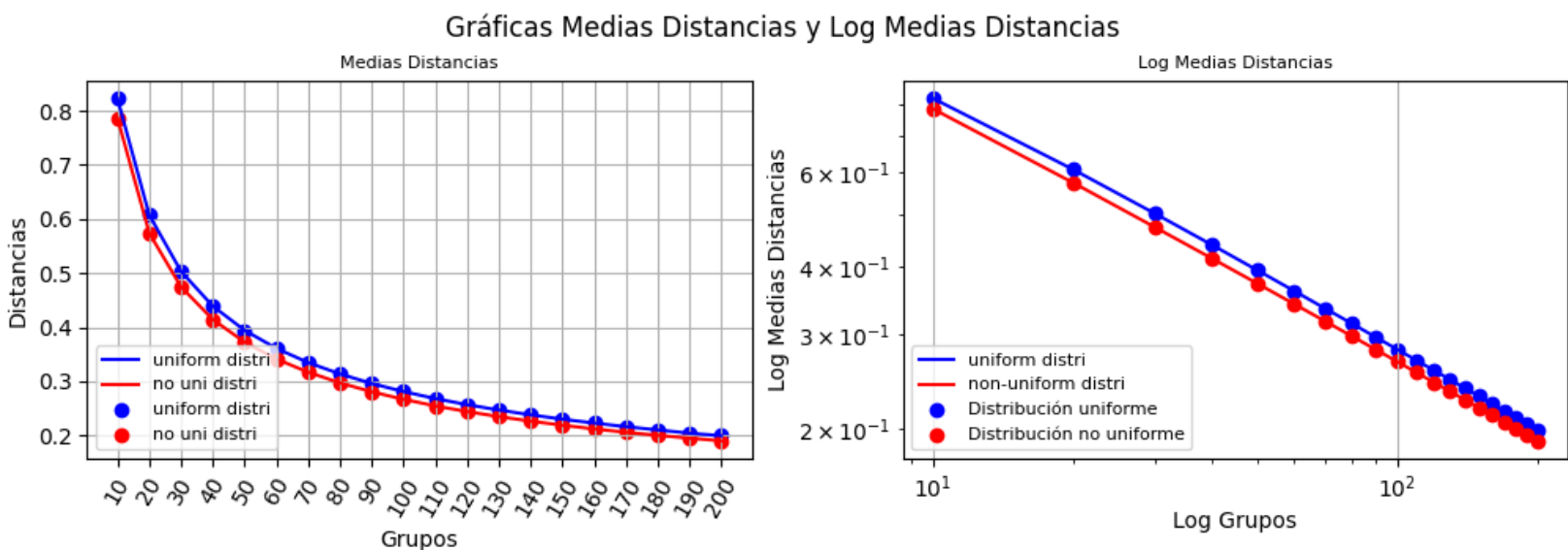
Cada grupo representaba grafos generados a partir de un número de nodos o puntos generados fijo, estos 20 grupos tenían valores de múltiplos de 10.

0.1 Resultados fase 1

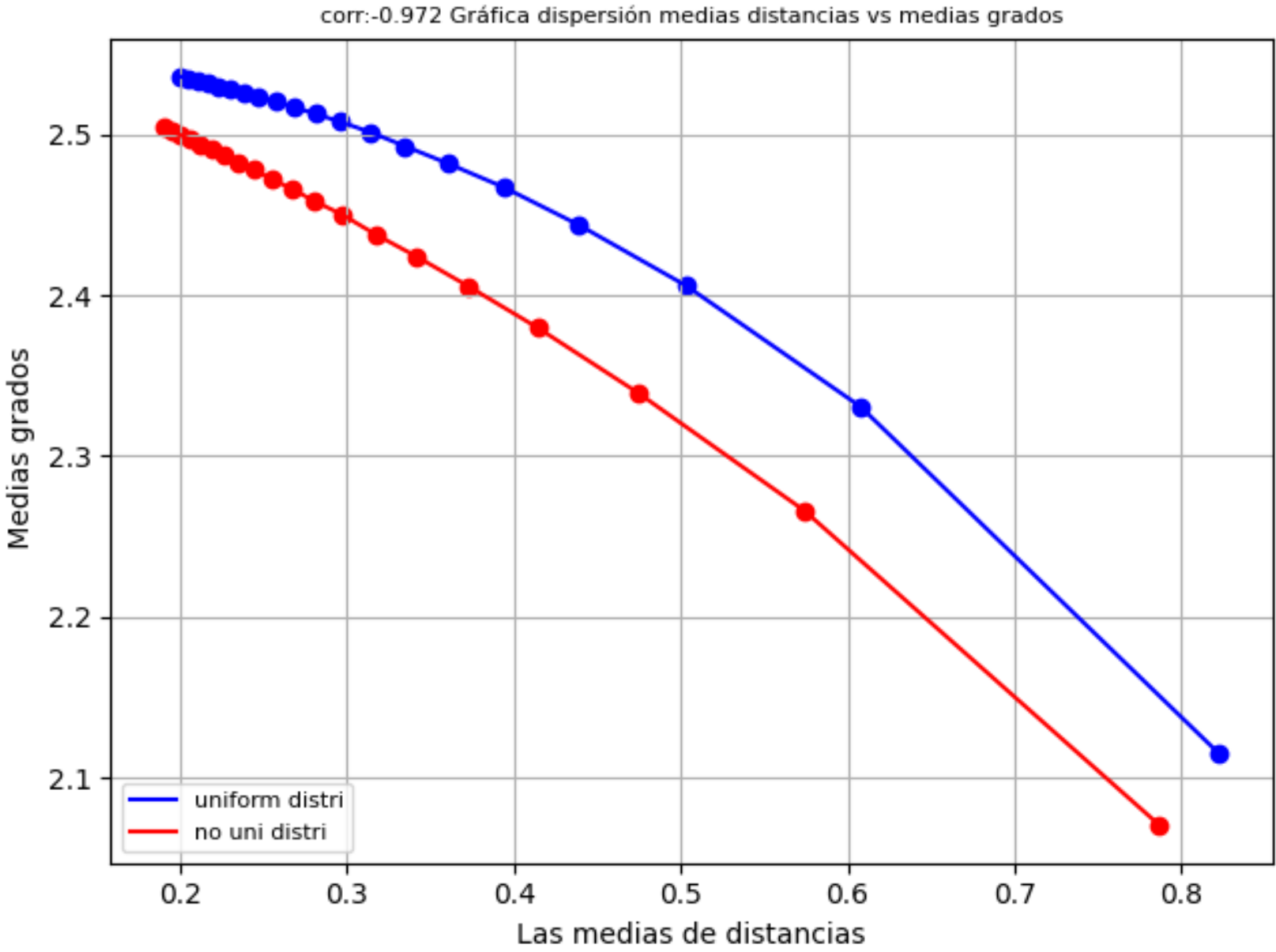
Recolectamos resultados de distancias , grados de los grafos y valores de medidas de tendencia central fidedignos para cada uno de los 40 grupos que calculamos en total.



Sin importar la distribución fuera uni o no uniforme observamos que al aumentar el número de puntos el grado promedio converge a 2.5



Las Distribuciones obtenidas siguen una ley de potencia.



	Uniform Distribution	Non-Uniform Distribution
Pearson Correlations		
Mean	-0.972000	-0.992000
Mode	-0.963000	-0.993000
Median	-0.975000	-0.995000

Los resultados de las correlaciones son interesantes porque nos indica que la propiedad de los grafos RNG que la relación de la distancia con respecto al número de grados en inversamente proporcional.

Fase 2

0.2 ¿Qué estamos haciendo?

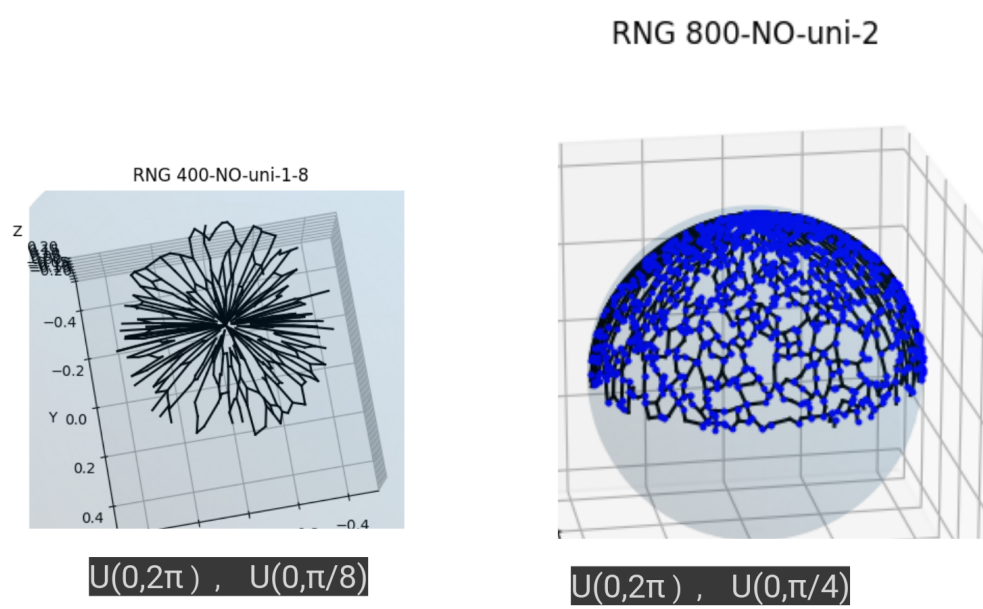
4. Usando las distribuciones encontradas en la fase 1, nos encontramos explorando distintas formas de construir una función de probabilidad,esta función asignará a cada par de puntos la probabilidad de que dos puntos tengan o no link.

La idea es que nuestra propuesta de algoritmo,construya las conexiones basado en un criterio probabilístico que decida la conectividad entre puntos.

Los resultados de nuestras funciones candidatas deberán minimizar el error de bondad entre los resultados del algoritmo probabilístico con respecto a los resultados del algoritmo original por un margen significativo, diferentes funciones como MSE,MAPE se consideran.

Fase 3

Analizar y entender el papel que tiene la curvatura de la esfera sobre nuestros resultados.



Nos encontramos reproduciendo las simulaciones realizadas pero restringiendo el espacio de distribución de puntos de manera gradual progresiva como en las figuras mostramos, hasta quedarnos solo con caparazones de longitudes reducidas.

Próximas líneas de investigación

El objetivo es proponer una nueva definición de entropía,la media de las áreas dentro de la cual cada punto puede moverse los vértices de gráficas RNG sin afectar su matriz de adyacencia,es decir,la región de espacio donde cada uno de los puntos podrían ubicarse sin modificar la configuración de la RNG original.

Utilizando el criterio de entropía que obtengamos podremos proponer un modelo estadístico mecánico que logre describir las propiedades macroscópicas de las RNG a partir del criterio de entropía.

References

- [1] Daniele Colosi and Pietro Dall'Olio. Relative neighbourhood graphs on the sphere. 2021.
- [2] Cole Foster, Berk Sevilmis, and Benjamin Kimia. *Generalized Relative Neighborhood Graph (GRNG) for Similarity Search*, pages 133–149. 09 2022.
- [3] Amaru Cuba Gyllensten and Magnus Sahlgren. Navigating the semantic horizon using relative neighborhood graphs, 2015.
- [4] Godfried T. Toussaint. The relative neighbourhood graph of a finite planar set. *Pattern Recognition*, 12(4):261–268, 1980.
- [5] Richard Zemel and Miguel Carreira-Perpiñán. Proximity graphs for clustering and manifold learning. In L. Saul, Y. Weiss, and L. Bottou, editors, *Advances in Neural Information Processing Systems*, volume 17. MIT Press, 2004.