

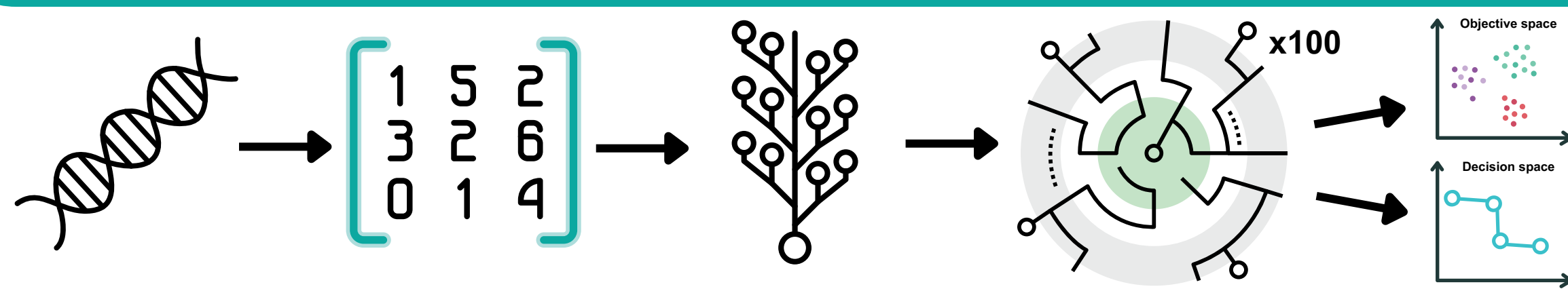
Relación entre los espacios de decisión y objetivo en el problema de inferencia de redes filogenéticas multiobjetivo

Israel Arias, Manuel Villalobos

israel.arias@usach.cl, manuel.villalobos@usach.cl

1.- Introducción

Las **redes filogenéticas** permiten representar relaciones evolutivas complejas que los árboles filogenéticos no pueden capturar. Estas redes se infieren (obtienen) mediante **optimización multiobjetivo**, la cual enfrenta retos al desconocerse la relación entre el **espacio de soluciones** (topologías de redes) y el **espacio objetivo** (criterios de optimización). Este desconocimiento podría llevar a la exclusión de soluciones relevantes, limitando la calidad de las inferencias filogenéticas generadas.



2.- Problema

Las estrategias de inferencia con optimización multiobjetivo se enfocan principalmente en el espacio objetivo, ignorando la diversidad topológica en el espacio de soluciones. Esta falta de atención a las diferentes topologías de las redes filogenéticas **puede estar causando sesgos en los resultados**, ya que redes diferentes pueden ser evaluadas de manera similar o redes parecidas ser puntuadas de forma distinta [1].

3.- Solución propuesta

A partir de diferentes conjuntos de datos genéticos, se aplican operadores de reordenamiento y mutación para generar redes filogenéticas. A estas redes se les calculan tanto métricas topológicas como criterios de optimización, como la verosimilitud y la parsimonia. Con esta información, se construyen el **espacio de solución** y el **espacio objetivo**, en los cuales, mediante técnicas de clustering, se agrupan las redes. Finalmente, se emplea el **índice de Jaccard** para **evaluar la correspondencia** entre ambos espacios.

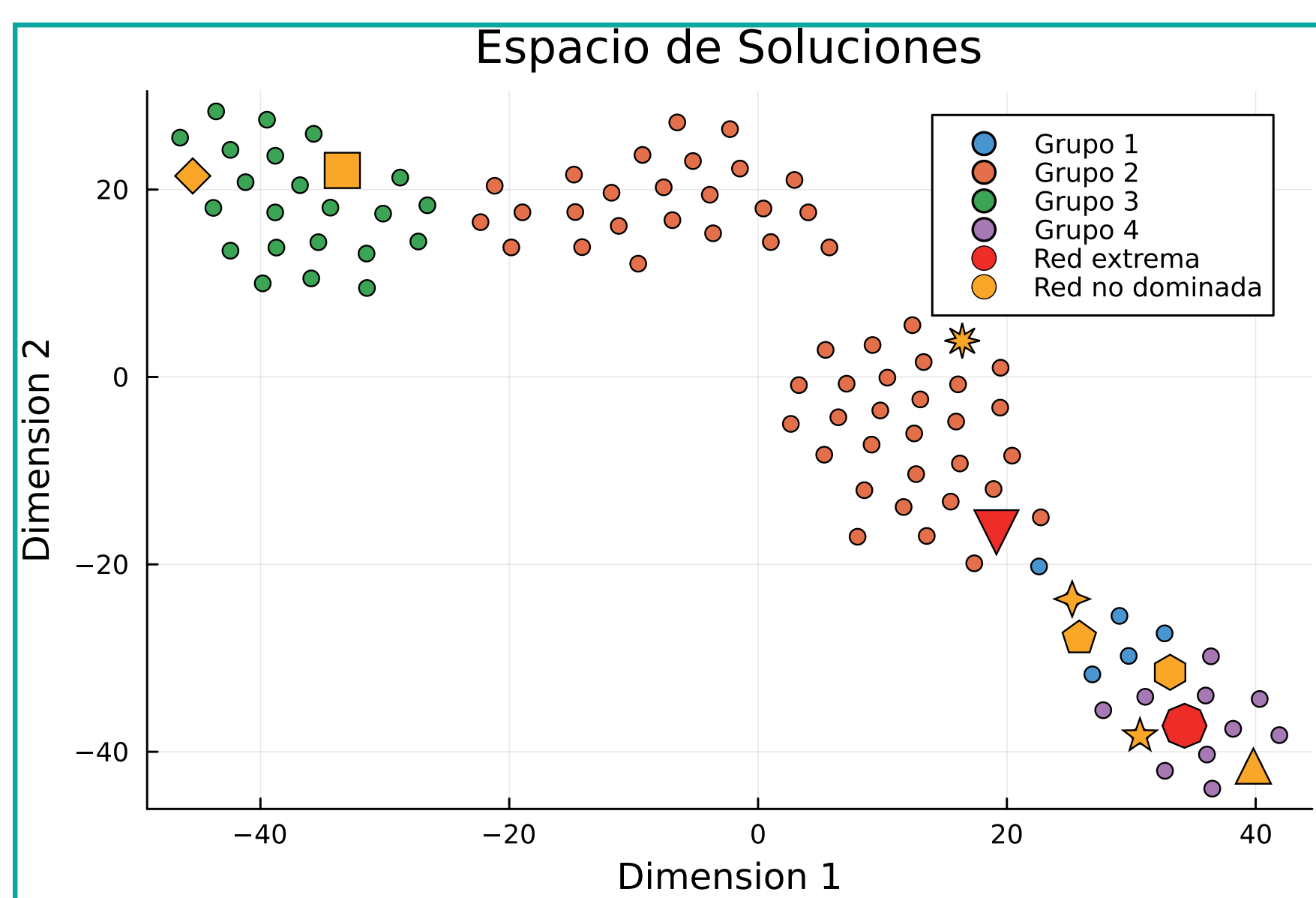
Data set	Pareto-set	PAR-LIK	
		Clusters	Jaccard Index
primates_14	10	4	0.26
rbcL_55	38	3	0.41
16s	20	3	0.47
HIV2_72	24	5	0.36
membracidae1_81	32	6	0.49

Average (AVE.)	0.39
Standard deviation	0.09

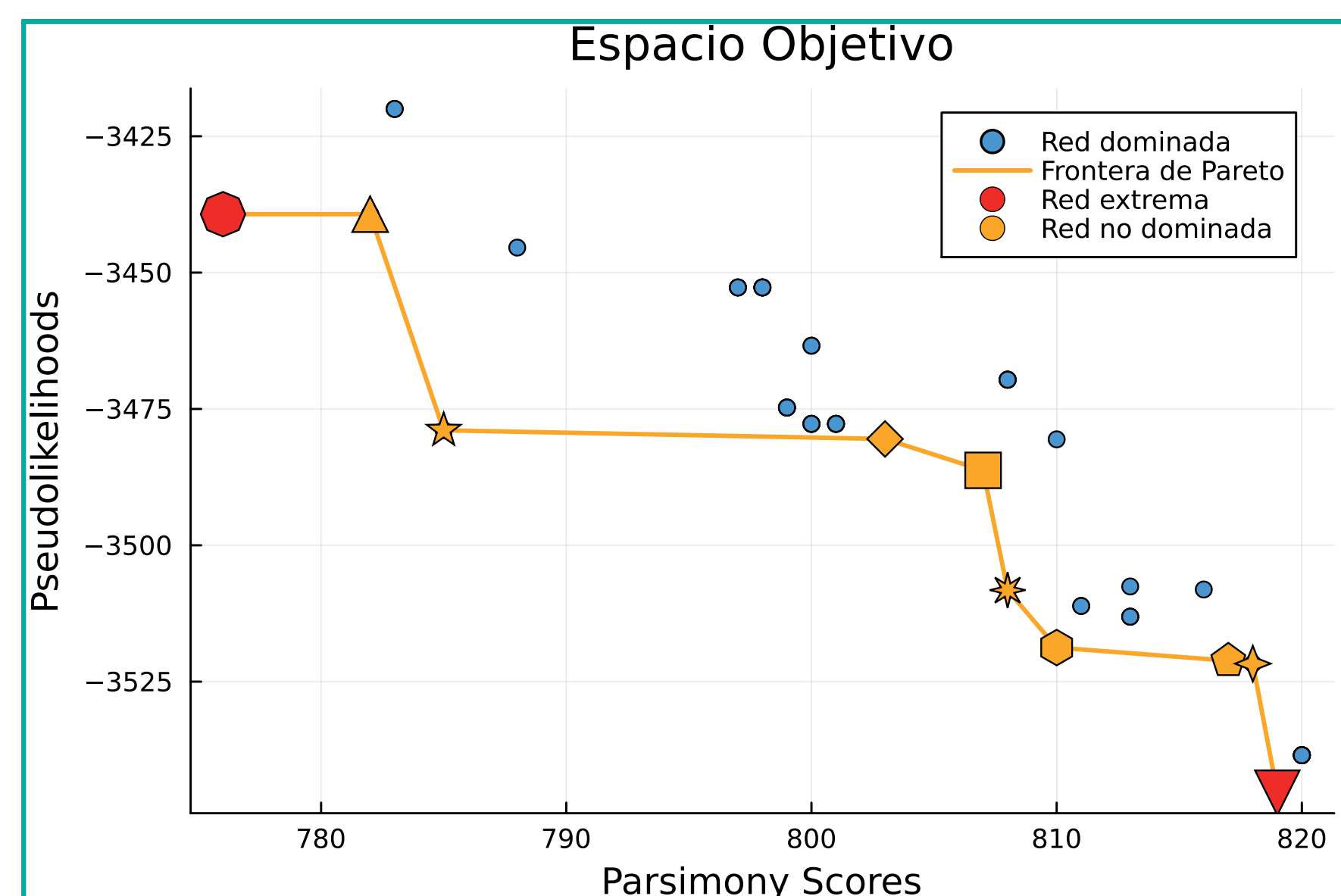
Comparación entre espacios de solución y espacio objetivo usando Índice de Jaccard

4.- Resultados

Los análisis revelan una **baja correlación** entre el espacio de soluciones y el espacio objetivo, con un **promedio de 0.39** en el **índice de Jaccard** para los set de datos utilizados. Redes con topologías distintas pueden tener **puntuaciones similares**, mientras que redes parecidas obtienen **resultados divergentes** en los criterios de optimización, como se observa en las **gráficas de ambos espacios**. Esto sugiere que las estrategias de optimización actuales podrían estar excluyendo soluciones valiosas al no considerar la diversidad topológica en el proceso de búsqueda, destacando así la **necesidad de desarrollar nuevos enfoques de optimización** que integren mejor ambos espacios y mejoren la calidad de las inferencias filogenéticas.



Espacio de soluciones - primates_14



Espacio objetivo - primates_14

5.- Referencias

[1] M. Villalobos-Cid, M. Dorn, y M. Inostroza-Ponta, "Understanding the Relationship Between Decision and Objective Space in the Multi-Objective Phylogenetic Inference Problem", en 2018 IEEE Congress on Evolutionary Computation (CEC), Río de Janeiro: IEEE, 2018, pp. 1-8. doi: 10.1109/CEC.2018.8477689.