

Optimización de la asociación de usuarios en un sistema Cell Free escalable

Yanela Varela Méndez, Gustavo González, Fernando Gregorio

Instituto de Investigaciones en Ingeniería Eléctrica «Alfredo Desages» (IIIE),
UNS-CONICET. Departamento de Ingeniería Eléctrica y de Computadoras (DIEC),
Universidad Nacional del Sur.
Bahía Blanca, Argentina
varelamendez@iiie-conicet.gob.ar

Resumen El aumento en la tasa de transferencia en las redes celulares ha sido propiciado por un aumento en la densificación, es decir, el incremento de puntos de acceso (AP) por unidad de área. Sin embargo, esta estrategia está dejando de ser eficaz, ya que también aumenta la interferencia de celdas aledañas. Recientemente se propuso un sistema basado en una estructura no-celular, llamado comunicaciones cell-free (CF), en el que un gran número de APs sirven de manera conjunta a un gran número de usuarios. Si bien este sistema es capaz de manejar la interferencia eficientemente, no es escalable, dado que el tráfico total por el fronthaul crece en función del número de usuarios. Esto motiva el desarrollo de una nueva arquitectura denominada CF escalable, en la que un subconjunto de los AP proporciona servicio a un subconjunto de los usuarios, limitando el tráfico por el fronthaul. En este trabajo se propone utilizar un algoritmo genético (GA) para hallar la asociación óptima entre usuarios y APs, maximizando la eficiencia espectral total del sistema. La escalabilidad se garantiza al limitar el número máximo de usuarios que cada AP puede soportar y el número máximo de APs al que un usuario puede asociarse. Un simple modelo de propagación en el espacio libre es usado para modelar el canal de comunicaciones. Los resultados de simulación muestran que la asociación eficiente de usuarios permite aumentar la eficiencia espectral al aumentar la potencia recibida, al mismo tiempo que se reduce la interferencia.

Palabras claves: Optimización con restricciones · Asociación de usuarios · Algoritmo Genético · Cell Free escalable.

1. Introducción

La quinta generación de redes celulares, conocida como 5G, ha sido un avance significativo en la conectividad inalámbrica, ofreciendo altas tasas de transmisión para el envío de datos (Banda ancha móvil mejorada, eMBB), ultra capacidad de respuesta, comunicaciones ultra confiables de baja latencia (URLLC) y capacidades de comunicación masiva de tipo máquina MTC/M2M o IoT [1].

Sin embargo, se prevé que, para los próximos años, las demandas de datos de una sociedad completamente digitalizada superarán las capacidades de la tecnología 5G. La siguiente fase evolutiva, conocida como 6G, plantea varios objetivos que cumplir: ofrecer más velocidad de transmisión de datos (se estima que hasta diez veces más) y bajar la latencia al máximo. Otros aspectos ya conocidos del 5G también se mejorarían en su sucesor: menor consumo de energía y mayor cantidad de dispositivos conectados de manera simultánea. Los desafíos significativos estarían en términos de eficiencia espectral y asignación de recursos [2].

El aumento del ancho de banda de transmisión en las redes celulares convencionales fue posible gracias al aumento en la densificación. La densificación de la red tiene que ver con la adición de más estaciones bases (BS) o puntos de acceso (AP) y la explotación de la reutilización espacial del espectro, mejorando así la capacidad de la red [3]. En la densificación, la potencia de la señal recibida por cada usuario aumenta, al igual que la interferencia de las celdas vecinas, lo que hace que la interferencia sea más difícil de manejar. Como resultado de este compromiso, la ganancia que se obtiene de la reducción del tamaño de la celda es limitada. La interferencia entre celdas es el principal problema de la densificación de la red. A medida que el tamaño de la celda disminuye, la relación señal-ruido más interferencia (SINR) queda dominada por la interferencia entre celdas. En este punto, seguir reduciendo el tamaño de la celda no incrementa la SINR, lo que implica que la densificación no puede proporcionar una mejora adicional del rendimiento. Además, si la celda continúa reduciendo su tamaño, el rendimiento podría deteriorarse debido a efectos de interferencia a pequeña escala [4].

MIMO Masivo (múltiples entradas y múltiples salidas, mMIMO) es otra tecnología utilizada en las redes celulares para manejar problemas relacionados con la interferencia entre celdas y la atenuación de canales en regiones distantes de la BS. En estos sistemas, el número de antenas es significativamente mayor que el número de usuarios a los que se debe prestar servicio. El uso de la tecnología mMIMO aumenta la velocidad de transmisión en el centro de la celda y también demuestra mejoras en los bordes de la celda. El uso de técnicas de formación de haces permite mejorar el rendimiento del sistema en comparación con las tecnologías MIMO convencionales. Sin embargo, a pesar de estos avances, todavía existen variaciones muy significativas en la cobertura y velocidad de transmisión dependiendo de la ubicación del usuario dentro de la celda [5].

Coordinated Multi-Point (CoMP) es otra tecnología que usa una red celular, que permite que múltiples APs cooperen y coordinen la transmisión y recepción para mejorar la calidad de la señal, mitigar la interferencia intercelular y optimizar la asignación de recursos de la red [6].

Al integrar los beneficios de las tres tecnologías descritas anteriormente: la capa física de Cellular Massive MIMO, el concepto de transmisión conjunta para puntos de acceso distribuidos en Coordinated Multi-Point (CoMP) y el régimen de despliegue de redes ultradensas, junto con una arquitectura donde no se distinguen celdas, se obtiene Cell Free (CF). A diferencia de CoMP, donde la coordinación ocurre entre APs dentro de celdas definidas, en un sistema CF no

hay límites de celda y todos los nodos de acceso cooperan de manera coherente. Por otro lado, como se ha mencionado, las redes ultradensas no son capaces de aumentar la SINR debido a la interferencia de celdas vecinas. Mientras que en un sistema CF los APs cuentan con información de todos los nodos, lo que permite un manejo óptimo de la interferencia entre usuarios. Por último, mientras que MIMO masivo es capaz de aumentar la SINR, sobre todo en usuarios cercanos al centro de la celda, un sistema CF va un paso más allá al utilizar un gran número de antenas distribuidas en el área de cobertura. Esto permite mejorar el servicio de todos los usuarios. En resumidas cuentas, la idea principal detrás de CF es eliminar las restricciones asociadas con las celdas alrededor de los APs, creando así una red dinámica y adaptable donde todos los nodos de acceso cooperan coherentemente para brindar servicios a los usuarios. Este enfoque innovador tiene sus raíces en la necesidad de proporcionar una cobertura uniforme y una mayor eficiencia espectral, especialmente en entornos con alta densidad de usuarios y alta demanda de datos [7].

En su concepción inicial, todos los APs del sistema CF daban servicio a todos los usuarios y el procesamiento de las señales se realiza de manera centralizada. Esto provoca un tráfico muy grande por el fronthaul, ya que las señales de todos los AP y las CSI deben transmitirse al nodo concentrador para su procesamiento. Dicho de otra manera, el sistema no es escalable. Para evitar este problema, surge CF escalable que consiste en permitir la asociación de cada UE a un subconjunto de los APs, que procesarán la información de manera local. Dicha asociación determina el desempeño del sistema ya que define la potencia recibida y el manejo de interferencia. La optimización de recursos en sistemas cell-free, tiene como problema central encontrar asociaciones óptimas de usuarios a los puntos de acceso que maximicen la eficiencia espectral global del sistema. Este problema desempeña un papel crucial en la mejora del rendimiento de la red y la experiencia del usuario, ya que una asignación eficiente de recursos puede garantizar un mejor uso del espectro disponible y una distribución equitativa de la capacidad entre los usuarios.

El trabajo se organiza de la siguiente manera: Primero, en la Sección II, se introducen las redes de CF como una evolución de las redes celulares convencionales para combatir la interferencia entre usuarios. En la Sección III, se presenta una implementación escalable de la arquitectura CF. En la Sección IV, se presenta el algoritmo genético utilizado para resolver el problema de la asociación óptima de usuarios a los APs, en un sistema CF-mMIMO. Luego, en la Sección V, se presentan los resultados de simulación obtenidos. Finalmente, en la Sección VI, se presentan las conclusiones del trabajo.

2. Sistema Cell Free

En cell free masivo MIMO (CF mMIMO), un gran número de APs equipados con una o múltiples antenas y distribuidos en una extensa área, sirven de manera conjunta a un gran número de usuarios (UEs), en el mismo recurso de tiempo y frecuencia. Los APs están conectados a través de enlaces fronthaul a unidades

centrales de procesamiento (CPU) responsables de la coordinación. Las CPUs están interconectadas mediante enlaces de backhaul. Las operaciones de precodificación y combinación se pueden realizar localmente en cada AP, centralmente en la CPU, o el procesamiento se puede distribuir entre el AP y la CPU. En la red MIMO convencional, una pequeña cantidad de APs se distribuyen geográficamente en una celda y sirven conjuntamente a todos los UEs, con información del estado de canal (CSI) global. Cada celda tiene una CPU para intercambiar el CSI local con otros. En la Fig. 1 se muestra un diagrama comparativo de las dos tecnologías.

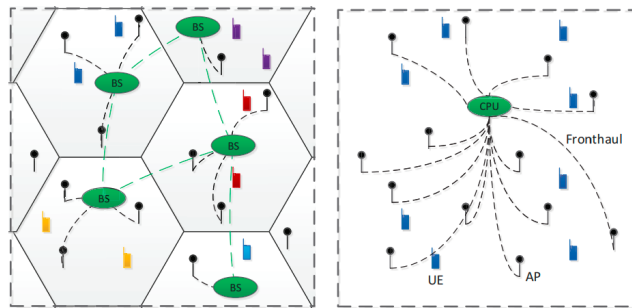


Figura 1: Comparación de la arquitectura de la red CF MIMO masivo y la red Sistemas MIMO convencionales.

En su concepción inicial, CF operaba con una única CPU, de la misma forma que la red MIMO. Es decir, un enfoque centralizado en el que la CPU es responsable de coordinar y procesar las señales de todos los UE. En este caso los APs son bastante sencillos ya que sólo necesitan recibir la señal y transferirla a la CPU central. Este enfoque demostró una eficiencia espectral (SE) muy alta [8]. Sin embargo, cuando se consideran redes con un gran número de UEs, la viabilidad de la implementación centralizada es cuestionable ya que la capacidad del fronthaul y de los APs debe crecer linealmente con la cantidad de usuarios [9]. En otras palabras, como es necesario enviar las señales transmitidas y recibidas por los APs y los CSIs al CPU central para el procesamiento, la forma original de Cell-Free Massive MIMO no es escalable. Bjornsson y Sanguinetti [10] sugirieron una arquitectura de red descentralizada en la que cada AP (más avanzado) estima localmente los canales de sus UEs asociados y utiliza esta información para procesar señales de datos localmente. Si bien utilizando este enfoque sub-óptimo la SE alcanzable es menor, se reduce considerablemente el uso del fronthaul y la red se vuelve escalable. En nuestro trabajo se utiliza el enfoque distribuido para la asociación de UEs a los APs y la gestión de restricciones. Esto permite que cada AP tome decisiones localmente sobre qué usuarios atender y cómo asignar los recursos. Por otro lado, el enfoque centralizado se utiliza para el cálculo de la eficiencia espectral.

3. Sistema Cell Free Escalable

En la Fig. 2 se muestra un ejemplo de una red CF escalable, donde la asociación de cada UE a un subconjunto de APs se denota con áreas de distinto color. Por otro lado, puede observarse que los APs pueden compartirse entre usuarios. Para evitar sobrecargar a los APs, es conveniente limitar el número de usuarios al que dan servicio. Debe notarse que cada AP solo tiene conocimiento del CSI de los UEs asociados, por lo que la señal de los otros usuarios va a resultar en interferencia que no se puede manejar. Por ese motivo, el enfoque CF escalable es subóptimo aunque de una complejidad de implementación mucho menor e independiente del número de UEs [11]. Por otro lado, es importante notar que si el área de cobertura es grande, se espera que los UEs se asocien a los APs cercanos. De esa manera la influencia de UEs o APs lejanos sería despreciable, lo que indicaría una baja pérdida de desempeño.

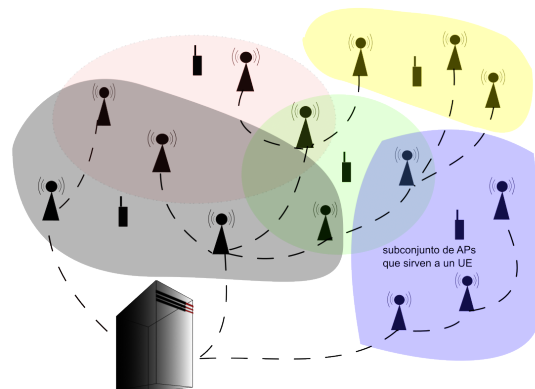


Figura 2: Ejemplo de cómo diferentes subconjuntos de APs sirven a UE en un sistema MIMO masivo sin celdas.

Es evidente que la manera en la que se asocian los UEs a los APs tiene un gran impacto en el desempeño del sistema, ya que determina la potencia recibida y el manejo de la interferencia. En la siguiente sección describiremos la técnica de optimización utilizada para encontrar la asociación entre UEs y APs óptima que maximiza la eficiencia espectral.

4. Técnica de optimización

Un algoritmo genético es un método de optimización basado en la selección natural de las especies. En esta técnica, un número inicial de individuos forman la población inicial que tendrá un cierto ajuste con la solución del problema.

Se utiliza una función basada en el problema para evaluar la calidad de las soluciones propuestas, llamada *fitness*. De la población, se seleccionan los mejores

individuos para la reproducción, con la esperanza de que se consigan individuos que estén más cerca de la solución óptima en la generación futura.

En base a esto, se distinguen tres grupos de soluciones: de *élite*, de *cruza* y de *mutación*. El grupo de *élite* constituye a las mejores soluciones, por lo que permanece inalterado. El *cruza*, como su nombre lo indica, se combina para formar un nuevo conjunto de soluciones mejores a sus predecesoras. Se considera al actual grupo padres y al nuevo grupo hijos. El último grupo, de *mutación*, sufre alteraciones aleatorias en su valor.

Posteriormente, se interroga a la nueva *generación* si ha terminado. La nueva *generación* es la reagrupación de los *hijos*, las *mutaciones* y la *élite*. Si no ha terminado, vuelve a llamar a la función fitness hasta obtener cambios apreciables entre sus miembros. En caso contrario, se dice que el programa ha convergido a una solución. Funcionamiento del algoritmo genético (Fig. 3):

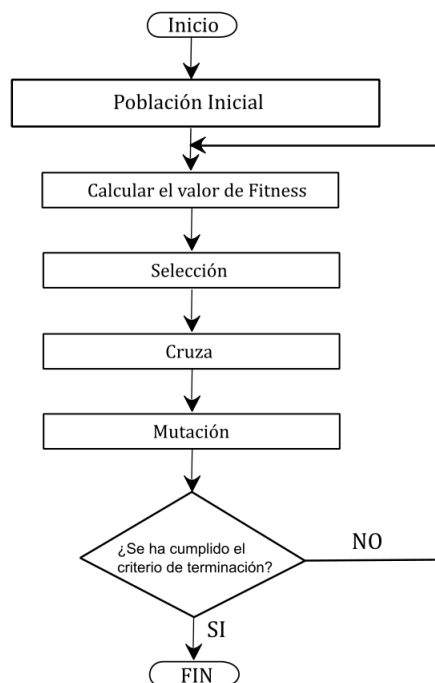


Figura 3: Diagrama de bloques de un algoritmo genético

4.1. Descripción de la implementación del algoritmo

El problema de optimización que se desea resolver consiste en maximizar la eficiencia espectral (SE), determinando la asociación óptima entre UEs y APs.

A fin de garantizar la escalabilidad del sistema, se agregan restricciones sobre el número máximo de APs a los que un UE puede conectarse y también el número máximo de UEs que un AP puede atender. El algoritmo genético utilizado, al estar basado en una búsqueda estocástica, tratará de hallar el óptimo global y no el punto crítico más cercano, como la mayoría de los algoritmos determinísticos.

La función de ajuste o función “fitness” en nuestro caso resulta la suma de las eficiencias espectrales individuales de todos los usuarios:

$$A = \sum_{i=1}^K SE_i \quad (1)$$

donde

$$SE_i = \sum_{j=1}^N x_{i,j} \cdot \log_2(1 + SNR_{i,j}) \quad \forall i \in [1, K] \quad (2)$$

la relación señal a ruido específica para el usuario i en función de su asignación al AP j se define como

$$SNR_{i,j} = \frac{P_i \cdot h_{i,j}}{\sum_{\substack{k=1 \\ k \neq i}}^K P_k \cdot h_{k,j} + R} \quad (3)$$

P_i es la potencia de transmisión del usuario i , $h_{i,j}$ es el canal entre el usuario i y el AP j , que describe la calidad de la señal en ese enlace, $\sum_{\substack{k=1 \\ k \neq i}}^K P_k \cdot h_{k,j}$ es la interferencia de otros usuarios sobre el AP j y R es densidad espectral de ruido. La variable booleana $x_{i,j}$ indica si el usuario i está conectado al punto de acceso j , teniendo en cuenta que $i \in [1, K]$ y $j \in [1, N]$. Finalmente, la función objetivo se expresa como $\max\{A\}$.

Para asegurar que la solución sea escalable, se establece que un AP puede conectarse a un máximo de S usuarios. En función de las variables de asociación, la restricción para el UE i puede expresarse como

$$\sum_{j=1}^N x_{i,j} \leq S \quad \forall i \in [1, K] \quad (4)$$

Por otro lado, la cantidad máxima de usuarios conectados simultáneamente a un AP j se limita, para evitar de esta manera la sobrecarga de los APs. Para expresar esta restricción, se utiliza la siguiente ecuación:

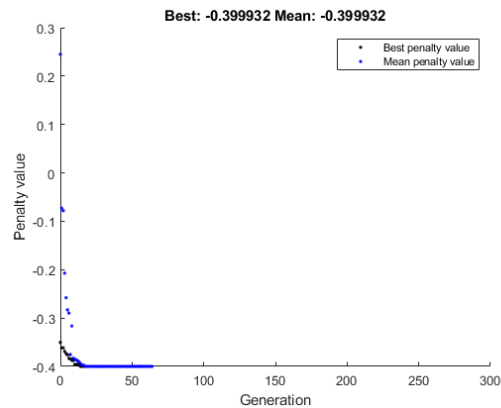
$$\sum_{i=1}^K x_{i,j} \leq L \quad \forall j \in [1, N] \quad (5)$$

donde L es el número máximo de usuarios conectados al AP.

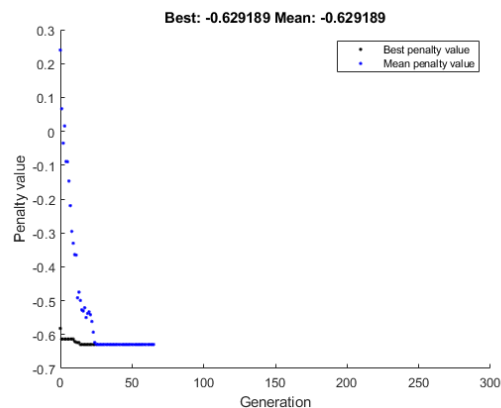
En la siguiente sección se presentará los resultados numéricos de la optimización.

5. Resultados

Se ha utilizado la función “ga” incluida en Matlab [12], que permite particularizar el algoritmo mediante un determinado número de opciones. El tamaño de la población es uno de los parámetros determinantes para el buen funcionamiento del algoritmo, para hacer la comparación se han probado diferentes tamaños. En la Fig. 4 se muestra el ajuste del mejor individuo y el de la media de la población para una población de 50 individuos y para una población de 300 individuos.



(a) Progreso del ajuste para una población de 50 individuos



(b) Progreso del ajuste para una población de 300 individuos

Figura 4: Progreso del ajuste para diferentes tamaños de la población

Se observa que el ajuste globalmente tiende a bajar, pero entre dos iteraciones puede suceder que se incremente su valor. Esto es debido a que la función de

ajuste se utiliza para determinar cuál es la mejor solución pero, además, las soluciones deben cumplir las restricciones. Por tanto, sucede que un individuo puede tener un mejor ajuste, pero no cumple bien las restricciones y por ello otro individuo con peor ajuste, pero que cumple mejor con las restricciones, es elegido como la mejor solución.

Se ha escogido una población de 300 individuos ya que, aunque todas las poblaciones lograron resultados similares, esta demostró ser más robusta que poblaciones más pequeñas.

En la Fig. 5 se observa la asociación eficiente de los usuarios a los APs, no solo por el hecho de que cada usuario se asocia al AP más cercano sino también porque cada AP solo puede soportar máximo 2 usuarios conectados, de los 3 usuarios que se generan aleatoriamente, evitando de esta manera la sobrecarga del sistema. Otro aspecto a destacar es que se logra la escalabilidad del sistema CF, ya que un subconjunto de 7 APs máximo pueden brindar servicio a cada usuario.

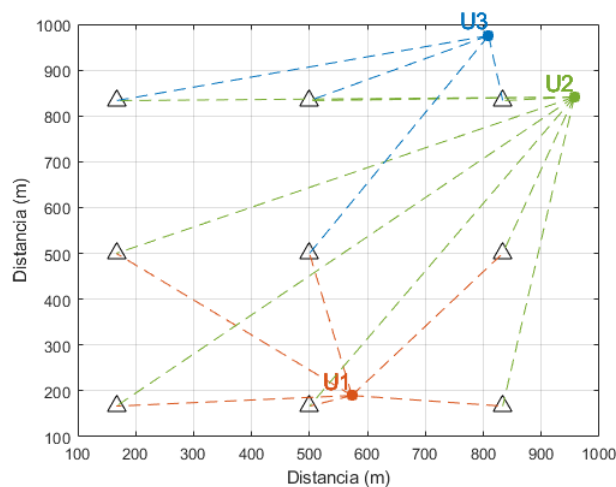


Figura 5: Asociación eficiente de los usuarios a los APs.

La Fig. 6 representa la eficiencia espectral óptima por usuario, donde el mejor resultado se alcanza para el U1 que se encuentra más alejado de los otros dos usuarios. Los usuarios (U2 y U3) están más cerca entre ellos, lo que aumenta la probabilidad de superposición de las señales transmitidas, y, por lo tanto, aumenta la interferencia mutua y disminuye la eficiencia espectral con respecto al U1. Si comparamos el valor alcanzado de la eficiencia espectral (SE) entre ambos usuarios (U2, U3), vemos que el U2 tiene mayor SE porque tiene más APs asociados, logrando una mayor diversidad de señales, lo que puede aumentar su potencia recibida y, por lo tanto, su SE.

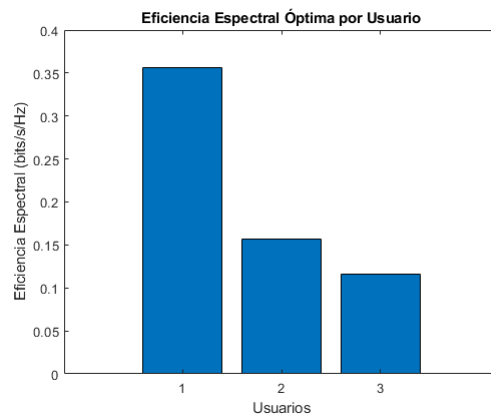


Figura 6: Eficiencia Espectral (SE) óptima para 3 usuarios.

Se utilizó un número pequeño de usuarios y de puntos de acceso para poder ilustrar gráficamente la asociación de los usuarios con los AP. En la Fig. 7 y Fig. 8 se presenta la SE óptima para un número mayor de usuarios, variando el número de APs. Los resultados obtenidos para la SE con diferentes valores de APs, son bastante similares pero se observa una ligera tendencia a la disminución de la SE, ya que al aumentar el número de AP disponibles, también aumenta la interferencia entre los puntos de acceso adyacentes. Además, se observa menor uniformidad entre los valores obtenidos cuando la cantidad de APs es mayor, esto puede ser porque algunos APs, están subutilizados.

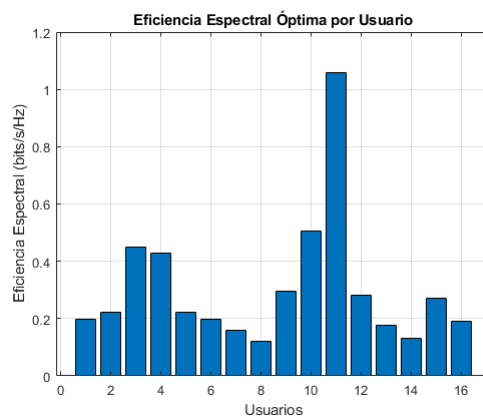


Figura 7: Eficiencia Espectral (SE) óptima para 16 usuarios.

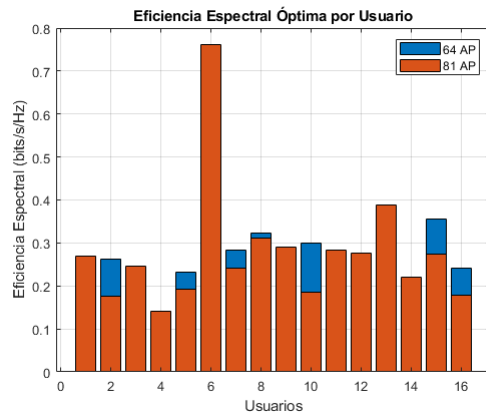


Figura 8: Eficiencia Espectral (SE) óptima para 16 usuarios.

Una mejor visualización del comportamiento de la eficiencia espectral, se tiene en la Fig. 9, donde se compara la SE promedio variando el número de APs en diferentes escenarios. La línea de color azul, representa el escenario escalable y muestra que la SE promedio disminuye para un número cada vez mayor de APs, debido a que la interferencia es mayor y algunos AP son subutilizados. La línea de color rojo, representa la SE promedio de un escenario centralizado. En este caso donde todos los UE se conectan a todos los APs la SE promedio no solo obtiene mejores resultados en comparación con el escenario escalable sino que además tiende a aumentar.

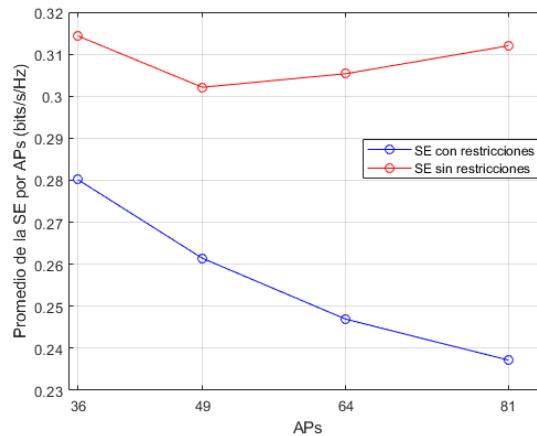


Figura 9: Comparación de la Eficiencia Espectral (SE) en un escenario escalable (azul) y centralizado (rojo).

Sin embargo, esta implementación no es realizable en un escenario real cuando la capacidad del fronthaul y de los APs crece linealmente con la cantidad de usuarios y es considerada como una cota superior de desempeño.

6. Conclusiones

En este trabajo se utilizó el algoritmo genético para hallar la asociación eficiente entre los usuarios y los APs maximizando la eficiencia total del sistema. Demostramos que los usuarios se asocian a los APs más cercanos, distribuidos de manera que no se sobrecargue el sistema, y que este a su vez sea escalable, ya que un subconjunto de APs, brindan servicio a los usuarios. En términos de eficiencia espectral, demostramos que la asociación eficiente aumenta la SE, y disminuye la interferencia. En el trabajo se incluyó el desempeño de un sistema centralizado (no escalable) que es utilizado a modo de comparación.

Referencias

1. Wong, Vincent Ws; Schober, Robert; NG, Derrick Wing Kwan (ed.). Key technologies for 5G wireless systems. Cambridge university press, 2017.
2. M. Matthaiou, O. Yurduseven, H. Q. Ngo, D. Morales-Jimenez, S. L. Cotton, and V. F. Fusco, "The road to 6G: Ten physical layer challenges for communications engineers," *IEEE Commun. Mag.*, vol. 59, no. 1, pp. 64–69, Jan. 2021.
3. N. Bhushan, et al. "Network densification: the dominant theme for wireless evolution into 5G." *IEEE Communications Magazine* 52.2, pp. 82-89, 2014.
4. Jeffrey G. Andrews, Xinchun Zhang, Gregory D. Durgin, and Abhishek K. Gupta, "Are we approaching the fundamental limits of wireless network densification?," *IEEE Communications Magazine*, vol. 54, no. 10, pp. 184–190, 2016.
5. L. Sanguinetti, E. Bjornson, and J. Hoydis, "Towards massive MIMO 2.0: Understanding spatial correlation, interference suppression, and pilot contamination," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 232–257, Jan. 2020.
6. R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H. Mayer, L. Thiele, and V. Jungnickel, "Coordinated multipoint: Concepts, performance, and field trial results," *IEEE Commun. Mag.*, vol. 49, no. 2, pp. 102–111, Feb. 2011.
7. Ngo, Hien Quoc, et al. Cell-free massive MIMO versus small cells. *IEEE Transactions on Wireless Communications*, vol. 16, no 3, p. 1834-1850, 2017.
8. H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson and T. L. Marzetta, "Cell-Free Massive MIMO versus Small Cells", *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1–17, 2017.
9. Gesbert, David, et al. Multi-cell MIMO cooperative networks: A new look at interference. *IEEE journal on selected areas in communications*, 2010, vol. 28, no 9, p. 1380-1408.
10. Bjornson, Emil; Sanguinetti, Luca. Scalable cell-free massive MIMO systems. *IEEE Transactions on Communications*, vol. 68, no 7, p. 4247-4261, 2020.
11. Hoffmann, Kryszkiewicz, Paweł. Evaluation of User-Centric Cell-Free Massive Multiple-Input Multiple-Output Networks Considering Realistic Channels and Frontend Nonlinear Distortion. *Applied Sciences*, vol. 14, no 5, p. 1684, 2024.
12. Pohlheim H. Tutorial: Genetic and Evolutionary Algorithm Toolbox for use with Matlab Documentationfor: Genetic and Evolutionary Algorithm Toolbox for use with Matlab version: toolbox 1.95 documentation 1.95, August 1999