

Handover en Redes Celulares: Machine Learning

Viviana Párraga-Villamar
Departamento de Electrónica,
Telecomunicaciones y Redes de
Información (DETRI)
Escuela Politécnica Nacional
Quito- Ecuador
viviana.parragav@epn.edu.ec

Pablo Lupera-Morillo
Departamento de Electrónica,
Telecomunicaciones y Redes de
Información (DETRI)
Escuela Politécnica Nacional
Quito- Ecuador
pablo.lupera@epn.edu.ec

Felipe Grijalva
Colegio de Ciencias e Ingenierías
Universidad San Francisco de Quito USFQ
Quito- Ecuador
fgrijalva@usfq.edu.ec

Resumen—Este documento presenta los avances de investigación realizados en referencia al tema de estudio del programa doctoral que consiste en la mejora del proceso de handover en las redes celulares en base a técnicas de Machine Learning (ML). La revisión de literatura de este tema se desarrolló mediante un Systematic Literature Mappings (SLM) para encontrar artículos relevantes sobre la problemática. En este artículo de divulgación se presentan algunas incursiones realizadas hasta la actualidad en esta investigación, que parten de un conjunto de datos reales tomados por aplicaciones celulares sobre el comportamiento de la red, en especial cuando se producen handovers. Es así como hasta el momento se han realizado las siguientes etapas: adquisición de datos, creación de nuevos features, clusterización de datos para manejar las diferentes ubicaciones, identificación de fallas en el handover, aplicación de técnicas de Machine Learning para predecir el handover, aplicación de sistemas de recomendación y métodos de evaluación del handover.

Keywords—Machine Learning, handover, red celular, sistema recomendación.

I. INTRODUCCIÓN

Machine Learning es parte de la inteligencia Artificial que busca aprender automáticamente de los datos para extraer conocimiento. Identificar y explotar patrones ocultos en los datos de "entrenamiento" son los objetivos del Machine Learning. Los patrones aprendidos se utilizan para analizar datos desconocidos, de modo que se puedan agrupar o asignar a los grupos conocidos [1].

Las redes celulares móviles se han convertido en generadoras y portadoras de datos masivos. Actualmente, las investigaciones sobre el avance de estas tecnologías se orientan en mejorar el servicio prestado a sus subscriptores, debido a la gran cantidad de datos que esta industria maneja. Es así, como los autores recopilan datos de las celdas de las redes móviles y los dispositivos celulares conectados a ciertas celdas, mediante mediciones como: ancho de banda del enlace descendente, intensidad de la señal antes y después de la medición del ancho de banda, relación señal/ruido antes y después de la medición; latencia de la descarga; entre otros [2].

Para el análisis masivo de datos es necesario utilizar métodos y herramientas adecuadas. Algunos autores utilizan únicamente métodos estadísticos para seleccionar las características más importantes de los datos; pero, aunque los métodos estadísticos se han aplicado con éxito en diversos campos, estos métodos por sí solos tienen la tendencia a producir resultados erróneos [3]. Es así como el uso de ML

como método de análisis de datos masivos en este campo ha aumentado.

Un problema que los investigadores han tratado de resolver es la falla presentada al momento de realizar handover en las redes celulares. El objetivo principal se ha centrado en optimizar el handover y la asignación de energía en las conexiones activas para maximizar el rendimiento de la red celular [4].

El proceso de handover se realiza cuando la intensidad de la señal de un teléfono móvil se reduce por debajo del mínimo, que depende de la tecnología celular y factores de configuración, y no es suficiente para mantener la conexión con una Estación Base (BS) y necesita conectarse a otra. En algunos casos, el handover es erróneo debido a que la intensidad de señal es mínima y la BS a la que debe conectarse está saturada, existe interferencias o solapamiento de canales, generando caídas de llamadas o fallos de conexión, ya que debe buscar otra BS para conectarse. Además, puede ocurrir que la BS a la que se conecta la estación móvil no sea la que proporcione la mejor intensidad de señal y todo ello conlleva un desperdicio de recursos computacionales o una reducción de la calidad de servicio [3].

De esta manera, este trabajo de investigación busca identificar las características que producen el handover en redes celulares y así identificar cómo se podría optimizar este proceso utilizando técnicas de ML. Además, se pretende conocer los tipos de handover que se presentan en la red y las condiciones en las cuales se ejecutan dichos procesos con el fin de proponer soluciones que permitan mejorar el desempeño de la red.

II. METODOLOGÍA

El proceso investigativo de este trabajo requiere tener las bases sustentables del área de conocimiento sobre los estudios relacionados al uso de ML para mejorar el handover en redes celulares. De esta manera, inicialmente, se realiza un SLM como estrategia de búsqueda de estudios sobre el tema, donde se define el campo de estudio, especialmente lo concerniente a las técnicas a utilizar en la investigación [5].

Es pertinente mencionar que existen dos formas para disponer de datos reales de las redes celulares, una opción, a partir de los reportes de la infraestructura de los operadores celulares, y la segunda, en base a mediciones realizadas en los teléfonos móviles. Con respecto a lo anterior, en este trabajo de investigación de forma paralela al SLM se procede a la recopilación de datos reales en redes de la ciudad de Quito-Ecuador mediante el uso de aplicaciones celulares NETWORK Cell Info y NETMONITOR.

Algunos autores para optimizar el proceso de handover han recopilado desde la BS de servicio los informes que contienen mediciones de Reference Signal Received Power (RSRP) de las estaciones móviles para que un agente centralizado elija la acción de handover. Cada BS ejecuta un algoritmo que infiere si una estación móvil necesita ejecutar el handover. Por otra parte, en otro estudio, con datos reales de tráfico LTE (Long Term Evolution) de una empresa de telecomunicaciones, se extraen las características del funcionamiento de la red que permiten decidir cuándo incluir el handover; este estudio también se realizó con el propósito de predecir el tráfico 5G [6]. Mientras que para predecir el ancho de banda y handover en redes 5G o 4G/5G, se utilizan como datos, información capturada mediante aplicaciones celulares con parámetros como ancho de banda, vecinos LTE, Received Signal Strength Indicator (RSSI), Reference Signal Received Quality (RSRQ), velocidad del equipo de usuario, entre otros [7].

Provistos de un dataset de datos confiables se procede a someterlos a un análisis mediante diferentes técnicas de ML, que permitan, en primera instancia, analizar el comportamiento del handover, y posteriormente, mejorar el desempeño de dicho proceso en las redes celulares. En consecuencia, en este estudio se consideran los tres tipos de aprendizaje más utilizados para probar hipótesis, como son supervisado, no supervisado y de refuerzo. Como resultado de aplicar estos estudios, análisis y técnicas se han obtenidos algunos resultados en el proceso investigativo.

III. RESULTADOS

El análisis de los parámetros de radiofrecuencia de las redes celulares tomados mediante aplicaciones celulares y posteriormente adecuados mediante pre-procesamiento que incluyó limpieza de datos, normalización de variables y eliminación de datos atípicos permitió denotar ciertos comportamientos de los parámetros en ciertos escenarios, donde el servicio ofrecido no es adecuada para el usuario.

Uno escenario es el proceso de handover, en donde, se evidencia ciertos comportamientos de los datos que han sido modelados mediante ML en busca de encontrar modelos que aprendan y logren mejorar el proceso de handover y con esto el servicio que ofrecen las operadoras de red celular.

A. Creación de features

Los features son los datos de entrada a los modelos de ML para su posterior entrenamiento. Es así, como a partir de los parámetros recopilados mediante mediciones de aplicación celular se procede a crear nuevas variables que permiten orientar el modelado al handover. Los features que se han creado son distancia entre ubicaciones que calculada a partir de la posición (latitud y longitud) de cada una de las mediciones obtenidas; velocidad del equipo de usuario; dirección de movimiento que se calcula la dirección de movimiento como un ángulo respecto al norte, siendo 0° hacia el este y aumentando en sentido contrario a las agujas del reloj.

También se calculó el rating en busca de definir una variable que valore la calidad de la señal recibida en cada una de las mediciones con las variables que presentaban mayor variabilidad. De la misma manera, Handover que es un feature de gran importancia dentro de este proceso investigativo, ya que permite filtrar los datos relacionados con el mismo. Conociendo que un proceso de handover se da cuando un equipo de usuario cambia de BS y teniendo información del

id de la BS a la que se conecta en cada medición se procede a calcular cuando ocurren los procesos de handover.

B. Clusterización de datos

Las mediciones tomadas cada segundo, en algunos casos, presentaban demasiados datos dentro de una misma ubicación. Por esta razón se procedió a clusterizar las mismas de acuerdo con la latitud y longitud. Como método de clusterización se utilizó k-means y los datos agrupados en cada clúster fueron tomados por su media.

Se realizó prueba del mejor número de clúster estableciéndose el valor de 100, donde se visualizó ubicaciones más separadas como se observa en la Fig. 1.

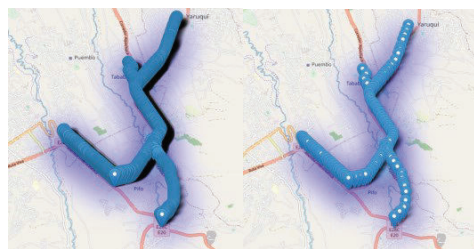


Figura 1. Visualización de antes y después de crear clústeres

C. Identificación de fallas en el handover

El proceso de ML requiere ser evaluado, pudiendo compararlo con un conjunto de datos que tenga un comportamiento correcto. Es así como se procedió a depurar la base datos original eliminando aquellas mediciones que podrían considerar un handover erróneo. Las condiciones tomadas en cuenta como una falla en el handover fueron: que luego de ejecutarse el handover la calidad de la señal no aumentó o la potencia disminuyó como debería ocurrir luego de un handover exitoso.

Además, se encuentra en análisis considerar el efecto Ping Pong en los handovers, esto es considerando el número de handovers por unidad de tiempo, la velocidad y la dirección del equipo de usuario.

D. Aplicación de técnicas de Machine Learning para predecir el handover

En el proceso de comprender el conocimiento de los datos recolectados se aplicaron diversas técnicas de ML para clasificar los datos con handover o sin éste.

Se inició realizando un proceso de resampling para balancear el número de registros con y sin handover; se dividió en un porcentaje de prueba y otro de entrenamiento; se escalaron todos los datos a un rango común para su interpretación; y se redujo la dimensionalidad mediante T-distributed Stochastic Neighbor Embedding (t-sne), que es un algoritmo para visualizar datos de alta dimensionalidad en un espacio de menor dimensión.

A continuación, se aplicaron varias técnicas como:

Regresión lineal entre las variables RSSI y Handover.

Regresión logística con varios grados del polinomio y lambdas para clasificar la existencia de handover.

KNN (K-Nearest-Neighbor), red neuronal y árbol de decisión (Fig. 2) para clasificar de acuerdo con la variable objetivo "handover".

Cada una de estas técnicas se sometieron a pruebas de evaluación.

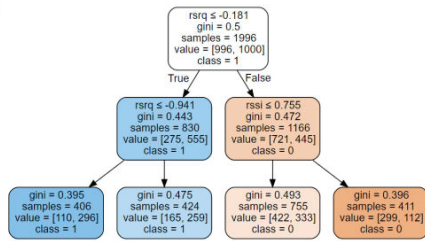


Figura 2. Árbol de decisión para predecir handover

E. Sistema de recomendación

Un sistema de recomendación es una técnica que pretende sugerir un conjunto de elementos, basándose en el historial de preferencias del usuario [9]. Se usó el filtrado colaborativo que recomienda ítems que el usuario no ha valorado en base a la búsqueda de patrones con valoraciones similares de otros usuarios.

Razón por la cual se buscaron hallazgos en este campo, considerando al sistema de red celular como un sistema donde las estaciones móviles (usuarios) reciben recomendaciones de a qué BS (ítem

) conectarse, en base a la variable rating (votación) calculada con anterioridad.

El filtrado colaborativo parte de una matriz de votaciones generada al relacionar, Radiobase (PSC PCI), estaciones móviles (clúster) y el rating. Luego se calcula la similaridad entre usuarios. A continuación, se define k usuarios similares mediante k-means, para estimar predicciones de BS no valoradas y finalmente recomendar los ítems con predicciones altas como en la Fig. 3.

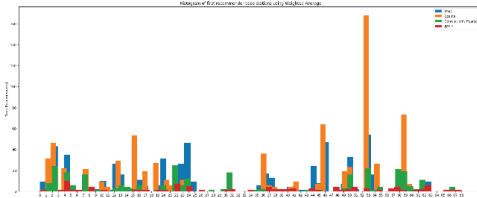


Figura 3. Histograma de las primeras estaciones base de recomendación utilizando la media ponderada

F. Métodos de evaluación del handover

Finalmente, se han considerado algunos métodos para la evaluación de los modelos creados con el propósito de validar sus respuestas. De manera general las técnicas revisadas fueron de clasificación y regresión, por tal razón se aplicaron métricas de rendimiento de acuerdo con la técnica utilizada en la creación del modelo como lo muestra la Tabla 1.

TABLA 1: MÉTRICAS DE EVALUACIÓN

Técnica ML	Métrica de evaluación
Regresión (Predicciones)	Error Medio Absoluto (MAE) Error Cuadrático Medio (MSE) raíz del MSE (RMSE)
Clasificación	Matriz de confusión Precisión Recall F1-Score Área bajo la curva de funcionamiento del receptor (ROC)

IV. DISCUSIÓN

De acuerdo con el estudio realizado hasta el momento, se definen tres enfoques esenciales para la correcta definición del problema caso de estudio. En primera instancia la fuente de datos debe recopilar datos en más intervalos de tiempo y con la ocurrencia de más handovers; segundo que la técnica de ML debe ser escogida cuidadosamente, ya que debe estar claro el objetivo del modelo y que features se requieren para conseguirlo. Y finalmente y más importante la evaluación requiere un conjunto de datos para validar la presencia de handover correctos.

V. CONCLUSIONES

Los datos reales requieren ser procesados antes de ser sometidos a técnicas de ML, siendo necesario crear nuevas variables como entrada al proceso de modelado. De esta manera se inició con 15587 mediciones, que luego de ser preprocesados quedaron en 13657, de donde 2563 presentaban handover de las cuales 239 fueron detectados como handover fallidos. Por otro lado, las técnicas de ML básicas no permiten obtener conocimiento relevante de la data, debe considerarse combinarlas para obtener mejores resultados, o usar técnicas como sistemas de recomendaciones. La regresión lineal presentó un MSE de 0.48, mientras que la regresión logística presentó un MSE máximo de 0.75 usando un lambda de uno y de grado 7. Por otro lado, KNN obtuvo una precisión de 0.65, a diferencia de red neuronal y árbol de decisión que obtuvieron precisión de 0.64 y 0.66 respectivamente.

REFERENCIAS

- [1] R. Boutaba *et al.*, "A comprehensive survey on machine learning for networking: evolution, applications and research opportunities," *J. Internet Serv. Appl.*, vol. 9, no. 1, p. 16, Dec. 2018, doi: 10.1186/s13174-018-0087-2.
- [2] M. Almaghrabi and G. Chetty, "A Deep Learning Based Collaborative Neural Network Framework for Recommender System," in *2018 International Conference on Machine Learning and Data Engineering (ICMLDE)*, Sydney, Australia: IEEE, Dec. 2018, pp. 121–127. doi: 10.1109/ICMLDE.2018.00031.
- [3] M. Azeem, M. Usman, and A. C. M. Fong, "A churn prediction model for prepaid customers in telecom using fuzzy classifiers," *Telecommun. Syst.*, vol. 66, no. 4, pp. 603–614, Dec. 2017, doi: 10.1007/s11235-017-0310-7.
- [4] D. Guo, L. Tang, X. Zhang, and Y.-C. Liang, "Joint Optimization of Handover Control and Power Allocation Based on Multi-Agent Deep Reinforcement Learning," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13124–13138, Nov. 2020, doi: 10.1109/TVT.2020.3020400.
- [5] J. M. Del Alamo, D. S. Guaman, B. García, and A. Diez, "A systematic mapping study on automated analysis of privacy policies," *Computing*, vol. 104, no. 9, pp. 2053–2076, Sep. 2022, doi: 10.1007/s00607-022-01076-3.
- [6] S. Zhao *et al.*, "Cellular Network Traffic Prediction Incorporating Handover: A Graph Convolutional Approach," in *2020 17th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, Como, Italy: IEEE, Jun. 2020, pp. 1–9. doi: 10.1109/SECON48991.2020.9158437.
- [7] L. Mei, J. Gou, Y. Cai, H. Cao, and Y. Liu, "Realtime Mobile Bandwidth and Handoff Predictions in 4G/5G Networks." arXiv, Apr. 26, 2021. Accessed: Feb. 15, 2023. [Online]. Available: <http://arxiv.org/abs/2104.12959>
- [8] V. P. Parraga, P. Lupera, and F. Grijalva, "Análisis estadístico general de mediciones de parámetros de la red celular en Quito," *Rev. Invest. En Tecnol. Inf.*, vol. 10, no. 21, pp. 120–131, 2022.
- [9] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey," *Knowl.-Based Syst.*, vol. 46, pp. 109–132, Jul. 2013, doi: 10.1016/j.knsys.2013.03.012.